Structural Basis for Complex Formation between Human IRSp53 and the Translocated Intimin Receptor Tir of Enterohaemorrhagic *E. coli*

&

Structural Characterisation of Amorfrutins Bound to PPARγ

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Jens Christian de Groot

aus Wermelskirchen

1. Referent: Honorarprofessor Dr. Dirk W. Heinz

2. Referent: Professor Dr. Michael Steinert

eingereicht am: 12.12.2011

mündliche Prüfung (Disputation) am: 23.04.2012

Druckjahr 2012

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

de Groot, J.C., Schlüter, K., Carius, Y., Quedenau, C., Vingadassalom, D., Faix, J., Weiss, S.M., Reichelt, J., Standfuss-Gabisch, C., Lesser, C.F., Leong, J.M., Heinz, D.W., Büssow, K., Stradal, T.E.B. Structural Basis for Complex Formation between Human IRSp53 and the Translocated Intimin Receptor Tir of Enterohemorrhagic E. coli. Structure 19, 1294-1306 (2011).

Contents

Abbreviations

9(S)-HODE	(9S)-9-Hydroxyoctadeca-10,12-dienoic acid
A	Absorption
Å	Ångstrøm (0.1 nm)
ABBA	actin-bundling protein with BAIAP2 homology
Abl	Abelson murine leukemia viral oncogene homologue
A/E	Attaching-and-effacing
AF4	Asymmetric flow field flow fractionation
Arp2/3	Actin-related protein 2/3
BAIAP2	Brain-specific angiogenesis inhibitor 1-associated protein 2
BAR	Bin-Amphiphysin-Rvs
Bcr	breakpoint cluster region
CCP4	Collaborative Computational Project No. 4
Cdc42	Cell division control protein 42
Cif	Cycle inhibiting factor
CRIB	Cdc42/Rac interactive binding
Crk	Chicken tumor nr. 10 regulator of kinase
CRKL	Crk-like protein
DRLPA	Dentatorubral-pallidoluysian atrophy
CV	Column volume
Da	Dalton
DAEC	Diffusely adherent *E. coli*
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EAEC	Enteroaggregative *E. coli*

ECP	E. coli common pilus
EH	Eps15 homology
EHEC	Enterohaemorrhagic E. coli
EIEC	Enteroinvasive E. coli
EPEC	Enteropathogenic E. coli
Eps8	Epidermal growth factor receptor kinase substrate 8
Erk1	Extracellular signal-regulated kinase 1
Esp	E. coli secreted proteins
Gb3	glycosphingolipid globotriaosylceramide
GST	Glutathione S-transferase
Hck	Hematopoietic cell kinase
HCP	Haemorrhagic coli pilus
HIV	Human immunodeficiency virus
hPPARγ	human peroxisome proliferator-activated receptor gamma
IEC	Ion exchange chromatography
IMAC	Immobilised metal ion affinity chromatography
IMD	IRSp53-MIM homology domain
IPTG	Isopropyl-β-D-thiogalactoside
IRSp53	Insulin receptor substrate p53
IRTKS	Insulin receptor tyrosine kinase substrate
ITC	Isothermal titration calorimetry
LBD	ligand-binding domain
M	Molecular mass
Map	Mitochondrial-associated protein
mDia	Mammalian homologue of *Drosophila diaphanous*
Mena/VASP	Mammalian enabled/Vasodilator-stimulated phosphoprotein
MIM	Missing in metastasis
MWCO	Molecular weight cut off
N-CoR	Nuclear receptor co-repressor
Nef	Negative factor

NleA	Non-LEE-encoded effector A (also known as EspI)
nTZDpa	Non-thiazolidinedione selective partial agonist
NMEC	Neonatal meningitis *E. coli*
N-WASP	Neural Wiskott-Aldrich syndrome protein
PAGE	Polyacrylamid gel electrophoresis
PDB	Protein Data Bank
PDZ	Postsynaptic density-95(PSD-95)-discs large-zone occludens-1
PGC-1α	Peroxisome proliferator-activated receptor gamma co-activator-1 alpha
PGC-2	Peroxisome proliferator-activated receptor gamma co-activator-2
15d-PGJ2	15-Deoxy-Delta-12,14-prostaglandin J2
$PI(4,5)P_2$	Phosphatidylinositol-4,5-bisphosphate
PPAR	Peroxisome proliferator-activated receptor
PTB	Phosphotyrosine binding
r.m.s.d.	Root mean square deviation
rpm	Rounds per minute
RXRα	Retinoid X receptor alpha
SDS	Sodium dodecyl sulphate
SH3	Src homology 3
Shank-1	SH3 and multiple ankyrin repeat domains 1
SMRT	Silencing mediator co-repressor for retinoid and thyroid-hormone
SPPARγM	Selective PPARγ modulators
Src	Proto-oncogene tyrosine-protein kinase Src (sarcoma)
SRC1	Steroid receptor co-activator-1
Stx	Shiga toxin
T3SS	Type three secretion system
TCEP	Tris(2-carboxyethyl)phosphine
Tir	Translocated intimin receptor
TirC	Carboxy-terminal cytoplasmic domain of Tir

TJ	Tight junctions
TLS	Translation/libration/screw
T_m	Melting temperature
TZD	Thiazolidinedione
UPEC	Uropathogenic *E. coli*
WASP	Wiskott-Aldrich syndrome protein
WAVE	WASP-family verpolin-homologous protein

Zusammenfassung

Strukturelle Grundlagen zur Komplexbildung von humanem IRSp53 mit dem *translocated intimin receptor* Tir enterohämorrhagischer E. coli

Enterohämorrhagische E. coli (EHEC) adhärieren an ihre Wirtszelle und induzieren lokale Aktinpolymerisation durch Translokation der Effektorproteine Tir und $EspF_U$, welche intrazellulär mit Proteinen der humanen IRSp53-Familie und N-WASP interagieren. In dieser Studie wurde die Komplexstruktur der aminoterminalen Domäne von IRSp53 mit einem Tir-Peptid gelöst. Es wurde gezeigt, dass zwei parallel orientierte Tir Moleküle an ein IMD-Homodimer binden. Durch diese Interaktionen vernetzt sich das Bakterium mit der Aktinpolymerisationsmaschinerie der Wirtszelle. Die Struktur zeigt eine bislang unbekannte, spezifische Peptid-Bindungsstelle auf der Oberfläche von IMD, welche zwischen IRSp53 und dem homologen IRTKS konserviert ist. Das Asn-Pro-Tyr (NPY) Motiv des Tir, essentiell für die Ausbildung von Pseudopodien-ähnlichen Strukturen, wird hierbei spezifisch von dieser Bindungsstelle erkannt. Die Verifizierung der Interaktion erfolgte durch gezielte Mutagenese und *in vivo* Bindungsstudien. Vermutlich nutzt IRSp53 die NPY-Bindungsstelle aber auch für weitere bislang unbekannte Interaktionen mit wirtseigenen Proteinen.

Strukturelle Charakterisierung der Bindung von Amorfrutinen an den *peroxisome proliferator-activated receptor gamma*

Amorfrutine wurden kürzlich als neue, ungiftige Naturstoffklasse identifiziert, welche eine hohe Affinität zum *peroxisome proliferator-activated receptor gamma* (PPARγ) aufweisen; ein nukleärer Rezeptor der an der Regulation des Lipid- und Glukose-Metabolismuses beteilig ist. Rosiglitazone, ein PPARγ-Agonist, wird derzeitig zur Behandlung der Typ-II-Diabetes verwendet. Der Agonist steigert die Insulinsensitivität, führt allerdings auch zu Nebenwirkungen wie Gewichtszunahme. Amorfrutine erhöhen ebenfalls die Insulinsensitivität, verhindern aber eine übermäßige Fettspeicherung und normalisieren den Triglycerid-Metabolismus. Somit sind Amorfrutine eine vielversprechende Alternative zur Behandlung oder Prävention von Typ-II-Diabetes oder allgemein dem metabolischen Syndrom. In dieser Studie wird die Aktivierung von PPARγ durch die derzeitig wirksamsten Amorfrutine, Amorfrutin 1, 2 und B strukturell charakterisiert. Die hochauflösenden Strukturen belegen eine, im Gegensatz zu Rosiglitazone, Helix H12 unabhängige Bindung und Aktivierung von PPARγ. Der Vergleich mit Kristallstrukturen von PPARγ im Komplex mit anderen Liganden ordnet die Amorfrutine der Gruppe der partiellen PPARγ-Agonisten zu. Einige partielle PPARγ-Agonisten zeigten bereits *in vivo* die Eigenschaft, Insulinsensitivität zu induzieren ohne Nebenwirkungen wie Adipogenese hervorzurufen. Da Amorfrutine

ungiftige Naturstoffe sind, könnten diese nicht nur als alternative Medikation bei Typ-II-Diabetes ohne Nebenwirkungen dienen, sondern wären auch als Nahrungsergänzungsmittel zur Prävention solcher Krankheiten geeignet.

Summary

Structural Basis for Complex Formation between Human IRSp53 and the Translocated Intimin Receptor Tir of Enterohaemorrhagic *E. coli*

Actin assembly beneath enterohaemorrhagic *E. coli* (EHEC) attached to its host cell is triggered by the intracellular interaction of its translocated effector proteins Tir and $EspF_U$ with human IRSp53 family proteins and N-WASP. This study reveals the structure of the amino-terminal IMD domain of IRSp53 in complex with a peptide derived from the Tir sequence, in which the homodimeric IMD domain binds two Tir molecules aligned in parallel. This arrangement provides a protein scaffold linking the bacterium to the host cell's actin polymerisation machinery. The structure uncovers a novel, specific peptide binding site on the surface of the IMD, which is conserved between IRSp53 and its homologue IRTKS. The Tir Asn-Pro-Tyr (NPY) motif, which is essential for pedestal formation, is specifically recognised by this binding site. The site was confirmed by mutagenesis and *in vivo* binding assays. It is possible that IRSp53 utilises the NPY binding site for additional interactions with as yet unknown partners within the host cell.

Structural Characterisation of Amorfrutins Bound to the Peroxisome Proliferator-activated Receptor gamma

Amorfrutins were recently identified as a non-toxic class of natural products with high affinity to the peroxisome proliferator-activated receptor gamma (PPARγ), a nuclear receptor with regulatory functions in lipid and glucose metabolism. The PPARγ agonist rosiglitazone is currently used in treatment of type II diabetes. It increases insulin sensitivity but has adverse effects such as weight gain. Amorfrutins were also shown to increase insulin sensitivity and to normalise triglyceride metabolism, but do not enhance undesired fat storage. Amorfrutins therefore represent a very promising alternative for treatment or prevention of type II diabetes and the metabolic syndrome in general. This study reveals the structural basis of PPARγ activation by the currently most promising amorfrutins, amorfrutin 1, 2, and B. In contrast to rosiglitazone, the high resolution structures revealed that PPARγ binding and activation by amorfrutins are helix H12-independant. Comparison with complex structures of PPARγ with other ligands confirmed that the amorfrutins belong to the group of partial PPARγ agonists. Few partial agonists, like amorfrutins 1, were already shown to separate adverse effects like adipogenesis from the induced insulin sensitivity *in vivo*. Thus, the amorfrutins have not only the potential to work as alternative drugs to treat diseases like type II diabetes without the induction of adverse effects. These non-toxic naturally compounds may also act as a dietary supplement to prevent these diseases.

1 Chapter I

Structural Basis for Complex Formation between Human IRSp53 and the Translocated Intimin Receptor Tir of Enterohaemorrhagic *E. coli*

1.1 Introduction

1.1.1 *Escherichia coli* and the pathogenic deputies

The ubiquitous commensal bacterium *E. coli* requires only a combination of few mobile genetic elements for changing into a highly adapted pathovar, capable of causing a range of diseases from gastroenteritis to extraintestinal infections of the urinary tract, bloodstream, and central nervous system (Croxen and Finlay, 2010). These genetic elements are called pathogenicity islands (PAIs), encoding for an arsenal of virulence factors like toxins, colonization factors, and other virulence factors. PAIs are located on plasmids or integrated into the bacterial chromosome and are flanked by prophages, transposons, integrases, and transposases to facilitate their mobility (Kirsch *et al.*, 2004).

Annually hundreds of millions of people worldwide suffer from pathogenic *E. coli* infections, rationalizing extensive research to characterise the different pathovars. Eight pathovars have been well characterised and can be classified as either diarrhoeagenic or extraintestinal *(*Kaper *et al., 2004)*. The group of diarrhoeagenic *E. coli* include enterohaemorrhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enteroaggregative *E. coli* (EAEC) and diffusely adherent *E. coli* (DAEC). Uropathogenic *E. coli* (UPEC) and neonatal meningitis *E. coli* (NMEC) are classified as extraintestinal *E. coli* (Table 1.1-1), causing urinary tract infections and sepsis/meningitis, respectively. All these pathovars need to adhere for their pathogenicity to the host cells,

which they achieve by characteristic mechanisms of attaching and exploiting the host cells (Croxen and Finlay, 2010). Only EIEC invades the host cell for multiplication and dissemination (Kaper *et al.*, 2004).

Table 1.1-1: Sites of pathogenic *E. coli* colonization and classification into diarrhoeagenic and extraintestinal *E. coli*. EHEC: enterohaemorrhagic *E. coli*, EPEC: enteropathogenic *E. coli*, ETEC: enterotoxigenic *E. coli*, EIEC: enteroinvasive *E. coli*, EAEC: enteroaggregative *E. coli*, DAEC: diffusely adherent *E. coli*, UPEC: uropathogenic *E. coli*, NMEC: neonatal meningitis *E. coli*.

	Diarrhoeagenic *E. coli*						Extraintestinal *E. coli*	
	EHEC	EPEC	ETEC	EIEC	EAEC	DAEC	UPEC	NMEC
Brain								●
Bloodstream							●	●
Large bowel	●			●	●			
Small bowel		●	●		●	●		
Kidney							●	
Bladder							●	

1.1.2 Enterohaemorrhagic *E. coli* (EHEC) and the haemolytic-uremic syndrome (HUS)

Enterohaemorrhagic *E. coli* (EHEC) are food-born human pathogenic bacteria that cause severe bloody diarrhoea. They are capable of triggering the fatal haemolytic uremic syndrome (HUS), the main cause of acute renal failure in childhood and adolescence, accompanied by low platelet count and destruction of red blood cells. The course of disease (Figure 1.1-1) initiates with watery diarrhoea after a mean incubation period of 3 days post infection with EHEC, accompanied by cramping abdominal pain (Karch, 2001). During the next 2 to 3 days, in most patients watery diarrhoea turns bloody. Most patients recover spontaneously within 7 days (Mead and Griffin, 1998). However, about 15% of infected patients younger than 10 years develop HUS (Karch, 2001). The incidence of HUS is about 1.0/100 000 children below 15 years in Germany and Austria, with a mortality of about 3-5% (Mead and Griffin, 1998; Zimmerhackl *et al.*, 2002). Up to 40% of HUS patients exhibit long term sequelae 10 to 15 years after onset of the disease, such as chronic renal failure that leads to a lifetime dependence on dialysis, or other serious complications, such as persisting renal hypertension, proteinuria, diabetes mellitus, or neurological disorders (Karch, 2001; Mead and Griffin, 1998; Zimmerhackl *et al.*, 2002). HUS is the most severe complication of a potentially avoidable food-borne infection, as the bacteria are transmitted through contaminated food in most cases.

1.1.3 The EHEC serotype O157:H7 account for 90% of all HUS cases

The EHEC strains of the O157:H7 serotype, so-named because it expresses the 157th somatic 'O' and the 7th flagellar 'H' antigen according to the Kauffmann-White scheme (Institut Pasteur, 2007; Orskov and Orskov, 1992), are the best studied EHEC strains, because they are clinically the most relevant pathogenic *E. coli* in Europe, North America, and Japan (Nataro and Kaper, 1998). Strains of the O157:H7 serotype were the first strains denoted as enterohaemorrhagic *E. coli*, which are believed to account for over 90% of all cases of HUS in industrial countries (Siegler, 1995), rationalising for instance the routinely examination of all diarrhoeal stools for *E. coli* O157 in the UK (Mead and Griffin, 1998).

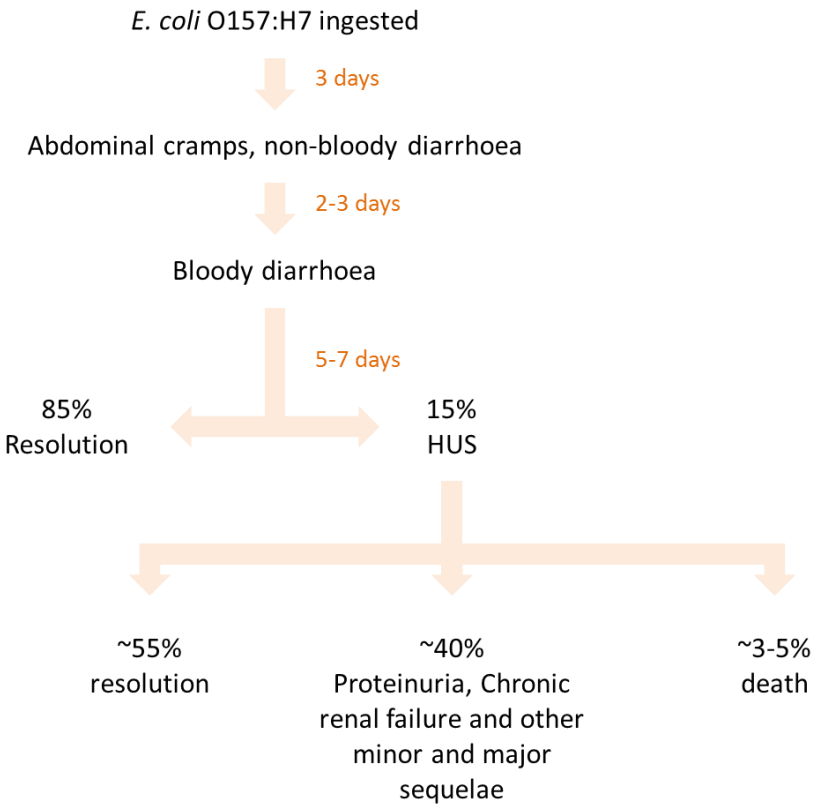

Figure 1.1-1: History of Infection with EHEC O157:H7.

Screening for EHEC O157:H7 infections relies on the inability of this serotype to ferment sorbitol rapidly, a phenotype distinct from most other *E. coli*, which leads to the formation of colourless colonies on sorbitol-containing MacConkey agar (Mead and Griffin, 1998). However, due to an increasing incidence of a sorbitol-fermenting serotype of EHEC O157:H7, as in the outbreak in Southern Germany of 2002 (Robert Koch-Institut, 2003), additional detection methods were introduced such as the use of antisera against the O157

antigen or assays for Shiga toxins, which are produced by EHEC (O'Brien *et al.*, 1984) but for instance not by the related EPEC (see also section 1.1.4).

The EHEC serotype O157:H7 possesses a highly conserved 92 kb plasmid, which encodes for different potential virulence factors. For instance, it encodes for EHEC haemolysin (Schmidt *et al.*, 1995) and a type II secretion system (Schmidt *et al.*, 1997). Although the plasmid encodes for several proteins which were expressed during the course of HUS, data supporting a role in the pathogenic process for each protein are still missing. Other virulence factors clearly play essential roles during infection, mainly the prophage encoded Shiga toxins that trigger the HUS (see section 1.1.4), but also other PAI encoded virulence factors, important for adhesion and dissemination of EHEC (see section 1.1.5.1).

1.1.4 Shiga toxins (Stxs) of EHEC cause HUS manifestation

Shiga toxins, or verotoxins, are the main virulence factors of EHEC and are the defining characteristics of Shiga toxin-producing *E. coli* (STEC), which were absent for instance in the EHEC related EPEC (Hayward *et al.*, 2006). STEC produce two major toxins with a sequence identity of 56%, Stx1 and Stx2 (Jackson *et al.*, 1987). Two variants of Stx1 (Stx1 and Stx1c) as well as several variants of Stx2 (Stx2, Stx2c, Stx2d, Stx2e and Stx2f) are known, with a sequence identity of 84-99% of the homologous to Stx2 (Johannes and Romer, 2010). The Shiga toxins are prophage encoded multi-subunit protein complexes that binds to the glycosphingolipid receptor Gb3 (CD77) on the surface of certain eukaryotic cell types, leading, depending on the target cell type, to apoptosis or the induction of an inflammatory response (Hayward *et al.*, 2006; Obrig, 2010). During the course of HUS in humans, Stx2 and the variant Stx2c are the two mainly produced Shiga toxins (Bockemühl *et al.*, 1997). HUS caused by these toxins typically affects the kidney and the gastrointestinal tract (Johannes and Romer, 2010).

1.1.4.1 Stxs and their causal role in HUS

The causal role of Stxs in HUS, summarized by Karch in 2010, has been inferred from: (1) Stx production is common to all HUS-associated *E. coli* isolate regardless of the serotype (Karmali *et al.*, 1985); (2) *in vivo* produced Stxs can be detected in faeces of HUS patients (Karmali *et al.*, 1985); (3) HUS patients develop immune responses to Stxs produced by the infecting strain (Karmali *et al.*, 1985); (4) intravenous injection of purified Stx causes severe disease in laboratory animals, which is prevented by Stx neutralizing antibodies (Richardson *et al.*, 1992); (5) in non-immune animals, circulating Stx binds rapidly to the vascular

endothelium of target organs (Richardson *et al.*, 1992); and (6) HUS-like disease has been described in racing greyhounds orally infected with *E. coli* O157:H7 (Hertzke *et al.*, 1995) and in baboons following injection of purified Stx (Taylor *et al.*, 1999).

1.1.4.2 Stxs are released into the gut by EHEC

Stx2 and its variant Stx2c seems to be more relevant than Stx1 during the course of HUS development and Stx2, in contrast to the almost completely cell associated Stx1, is found in the extracellular fraction (Shimizu *et al.*, 2009). Therefore, huge effort has been invested to examine especially the mechanism of the Stx2 release into the gut. To date, there are two mechanisms suggested to be involved in the release of Stx2 into the gut: (1) a specific induction of the Stx2-encoding prophage and (2) a specific Stx2 secretion mechanism (Shimizu *et al.*, 2009).

The release of Stx2 by EHEC into the gut as a result of specific induction of the Stx2-encoding prophage is mediated by the SOS response, a repair mechanism of bacteria induced by DNA damage according to different kind of triggers, like antibiotics or DNase colicins (Toshima *et al.*, 2007). The SOS response is regulated by LexA and the recombinase A (RecA) and involves the expression of more than 40 genes (Croxen and Finlay, 2010). Under normal condition, RecA is transcriptionally repressed by LexA. However, DNA damage activates RecA by the interaction with single stranded DNA, catalysing the cleavage of its repressor LexA, but also the cleavage of the protein CI, a repressor of the lysogenic lambda phage. This leads to an increased production of the prophage encoded Shiga toxins, replication of the lambda phage and finally the induction of the lambdoid phaged-mediated lysis of the host cell and thus the release of the toxins into the gut. Therefore, EHEC infection treatment with antibiotics is often counterproductive as it leads to the release of the intracellular toxins and even triggers their expression, as e.g. shown for mitomycin C that greatly induce Stx2 expression (Shimizu *et al.*, 2009; Wong *et al.*, 2000; Zhang *et al.*, 2000).

Along with the SOS response-mediated Stxs release by lysis of the host cell, a Stx2 specific secretion mechanism is also thought to contribute to the translocation of this toxin into the gut (Shimizu *et al.*, 2007; Shimizu *et al.*, 2009). Ser31 of the Stx's B subunit (see also section 1.1.4.3) seems to be recognised by a Stx2 secretion system in EHEC, as a S31N mutant results in an inhibited extracellular but an increasing cell-associated localisation of Stx2 (Shimizu *et al.*, 2007; Shimizu *et al.*, 2009). This suggested mechanism causes a continuously release of basally expressed Stx2, in contrast to the cell-associated Stx1

(Shimizu *et al.*, 2009). However, the detailed mechanism of Stx2 release into the gut remains still elusive.

1.1.4.3 Stxs trafficking

X-ray crystallography revealed a hexameric structure with an AB_5 molecular configuration of the Shiga toxin, with a molecular mass of 32 kDa of the A and 7.7 kDa of each B subunit, respectively (Fraser *et al.*, 1994; Stein *et al.*, 1992). The B subunits, encircling the carboxy-terminus of the A subunit, dictate the localisation of the Shiga toxins within the bacterial cells and thus are responsible for the almost complete cell associated localisation of Stx1 in EHEC and the generally extracellular localisation of Stx2 (Shimizu *et al.*, 2007; Weinstein *et al.*, 1989). Beside its crucial role regarding bacterial localisation, the B subunit also dictates the toxins internalization by the host cells. The B subunit exhibit two high affinity and one low affinity binding site for three Gb3 molecules, leading to a binding constant in the nanomolar range (Bast *et al.*, 1999; Fuchs *et al.*, 1986; Soltyk *et al.*, 2002). The interaction of the B subunit with Gb3 on the surface of paneth cells (Schuller *et al.*, 2007), located at the bottom of the crypt, induces endocytosis of the toxins. The toxins are then transported in a retrograde manner via the golgi to the endoplasmic reticulum (Johannes and Romer, 2010). Thereafter the A subunit is activated by an enzymatically cleavage, released into the cytoplasm and inactivates the eukaryotic ribosome by the removal of a single adenine base from the 28S rRNA within the large 60S ribosomal subunit (Endo *et al.*, 1988). This inhibits protein synthesis which may result in apoptosis (Obrig *et al.*, 1988).

However, it should be noted that Stx is shown to be also present in human Gb3-negative intestinal cells, possibly after being taken up by macropinocytosis (Malyukova *et al.*, 2009). Macropinocytosis is a non-selective uptake mechanism based on actin cytoskeleton rearrangement at the plasma membrane (Lim and Gleeson, 2011), as it occurs during the course of EHEC infection (see section 1.1.5.2). However, inside these cells, Stx does not prevent protein synthesis or induce apoptosis (Schüller *et al.*, 2004), but may inhibit chemokine expression and that in turn would inhibit inflammation (Gobert *et al.*, 2007). Furthermore, Stx2 increases the expression of the cell surface protein nucleolin in epithelial cells, shown to interact with the cell surface protein intimin of EHEC and thus may contribute to the bacterial adhesion (see also section 1.1.5.2) (Robinson *et al.*, 2006; Sinclair and O'Brien, 2002).

1.1.4.4 Stxs and the fatal consequences in patients

Stxs provoke local damage of the microcirculation, causing vasculitis, which exacerbates mucosal infraction and leads to bleeding into the bowel and thus bloody diarrhoea (Johannes and Romer, 2010), the primary clinical symptom of typical diarrhoea-associated HUS. Cattle are the main reservoir of STEC, but surprisingly without any clinical symptoms of infection, which may be explained by the lack of the Gb3 receptor (Pruimboom-Brees *et al.*, 2000).

The Stxs are also able to travel by the blood stream to other organs with fatal consequences. Peripheral blood monocytes, shown to be Stxs resistant despite their Gb3 receptor, respond to Stx interaction with releasing pro-inflammatory cytokines (Harrison *et al.*, 2004; Ramegowda and Tesh, 1996) that in turn stimulate Gb3 expression by endothelial cells of different organs (Harrison *et al.*, 2004; Johannes and Romer, 2010). Thereby, organs like the lung, the central nervous system, but most notably the kidney, were sensitized to Stx intoxication. The primary severe damage of the kidney may be explained by its most abundant amount of expressed Gb3, but also by the high volume of blood flow and the high blood filtration rate and thus a more excessive interaction with the toxin, compared with other organs (Obrig, 2010). As mentioned above, the most affected patients are younger than 10 years which may be ascribed to the fact that only pediatric kidneys express Gb3 in the glomeruli (Lingwood, 1994), cells located at the blood-urine barrier. The microenvironment of the Gb3 receptor seems to be also critical for the sensitivity of the receptor to the toxins, as Gb3 was shown to interact with membrane components like cholesterol, other glycolipids, fatty acids, or other proteins (Lingwood *et al.*, 2010a; Lingwood *et al.*, 2010b; Mahfoud *et al.*, 2010; Nutikka and Lingwood, 2004).

Furthermore, Stxs trigger a thrombotic mechanism. The toxin induces platelet-leukocyte aggregation and expression of tissue factor in the kidney, the brain (Sugatani *et al.*, 2000), and blood cells (Stahl *et al.*, 2009). Tissue factor containing microparticles were released, enter the circulation and initiate the conversion of fibrinogen to fibrin, resulting in the formation of an insoluble fibrin clot (Bach, 1988) and thus increasing the risk of vascular congestion.

Even if Shiga toxins are now commonly accepted as main virulence factors of pathogenic STEC (Karch, 2001; O'Brien *et al.*, 1992; Obrig, 2010), their role during the pathogenesis of the typical diarrhoea-associated HUS is not yet completely understood.

1.1.4.5 Non-Stx-HUS or atypical-HUS

The Non-Stx-HUS is less common than the typical Stx mediated and diarrhoea-associated HUS, is not diarrhoea-associated, accounts for only 5 to 10% of all cases of the disease and,

in contrast to the Stx mediated HUS, may manifest at all ages, but is more frequent in adults (Noris and Remuzzi, 2005). As the name already implies, the non-Stx-HUS is not caused by Stx, but by non-Stx producing bacteria (*Streptococcus pneumonieae*), or e.g. by viruses such as HIV, different drugs, or systemic diseases like lupus (Noris and Remuzzi, 2005). Therefore, this type of HUS is not EHEC mediated and only mentioned in this work for the sake of completeness.

1.1.5 The mechanism of EHEC to colonise the human intestine

After ingestion of the pathogenic bacteria, EHEC have to compete against the bacteria of the healthy intestine flora to colonise the gut. To accomplish this task, bacteria take advantage of different evolutionary evolved genetic modifications (see section 1.1.5.1), leading to expression of different effector proteins that induce host cell's actin cytoskeleton rearrangement and promote their tight adhesion and dissemination. Furthermore, they induce diarrhoea, which flushes out less adherent bacteria and thus helps EHEC to colonise the intestine (see section 1.1.5.2).

1.1.5.1 The locus of enterocyte effacement (LEE)

EPEC and most EHEC strains, including EHEC serotype O157:H7, contain almost identical copies of the locus of enterocyte effacement (LEE) (Perna *et al.*, 1998). The LEE is a chromosomal pathogenicity island that encodes for additional virulence factors aside from the prophage encoded Shiga toxins. These factors are essential for EHEC-mediated rearrangement of the host cell's actin cytoskeleton and facilitate adhesion and dissemination of the pathogens. The LEE is organized into five major operons, LEE 1 to 5. The LEE, the most important PAI, has evolved over the last 4.5 million years (Reid *et al.*, 2000) to a tightly packed sequence which contains only the genes of capital importance during the course of infection (Kirsch *et al.*, 2004). It comprises 41 open reading frames, encoding transcriptional regulators, a type III secretion system (T3SS), chaperones, translocators, the adhesin intimin, and effector proteins like the translocated intimin receptor (Tir) (Figure 1.1-2) (Garmendia *et al.*, 2005).

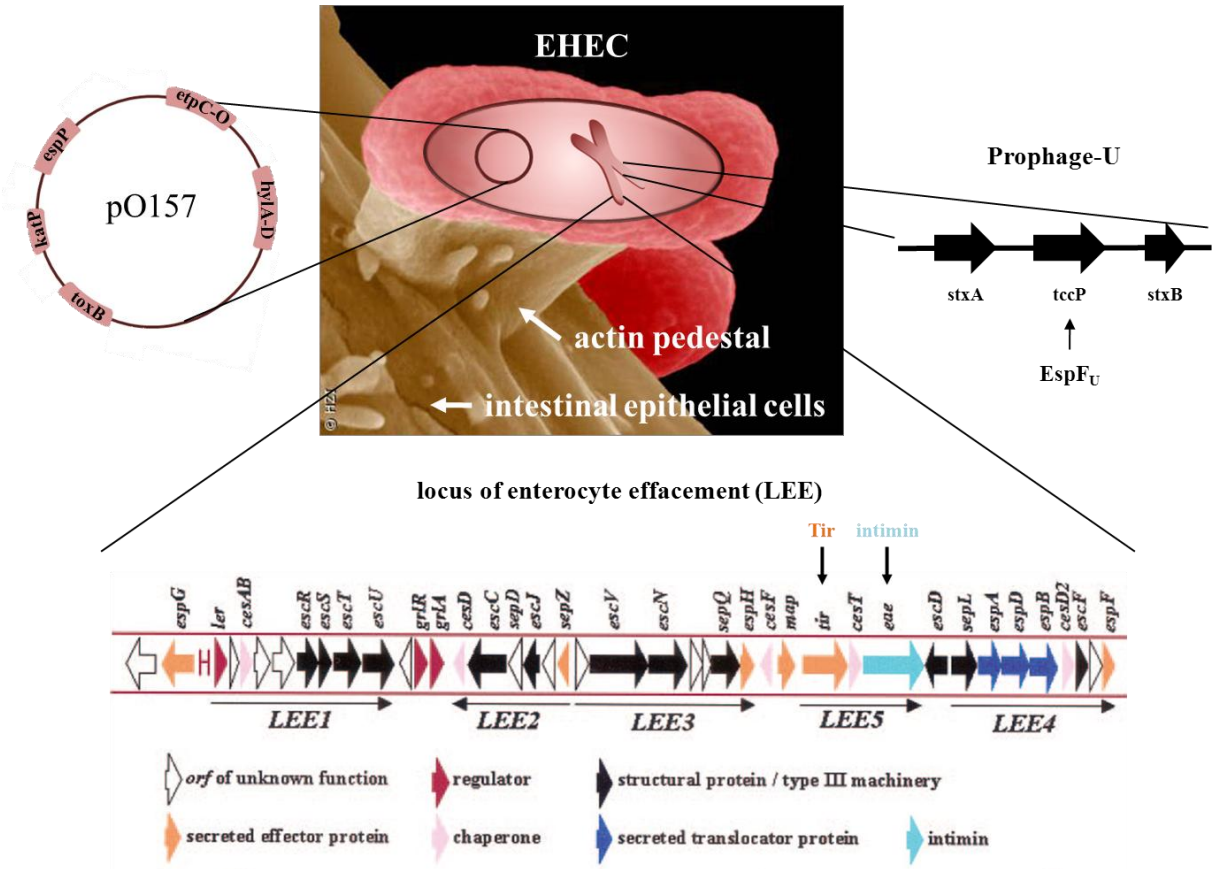

Figure 1.1-2: Localisation of the major genes involved in EHEC O157:H7 pathogenicity. Shown are the locus of enterocyte effacement (LEE), the $EspF_U$ and Shiga toxins encoding prophage-U, and the plasmid pO157, encoding for proteins like haemolysin (HlyA). The chromosomal pathogenicity island LEE of *E. coli* O157:H7 is shown according to Garmendia *et al.* (2005). The five putative polycistronic operons (LEE1 to LEE5) of EHEC LEE are indicated. EHEC image: Manfred Rohde, HZI.

EPEC as well as EHEC infections lead to so-called attaching-and-effacing (A/E) lesions, defined by the attachment of the bacteria to the epithelial cells of the intestinal mucosa and the subsequent loss (effacement) of their microvilli. EHEC require the LEE-encoded proteins intimin, the corresponding receptor Tir, and $EspF_U$ to induce A/E lesion formation in the host cell. The prohage-U encoded effector protein $EspF_U$, named according to the similarity to the LEE encoded effector protein EspF (Campellone *et al.*, 2004), and Tir need to be translocated by the T3SS into the host cell. These three proteins are sufficient to induce A/E lesions (see section 1.1.5.2) by a pathway related to EPEC infections (see section 1.1.5.3).

A/E lesions seem to play a crucial role during EHEC O157:H7 infection. However, LEE is not essential for the pathogenesis of all STEC, as some cases of severe disease, including HUS, are also caused by LEE-negative STEC strains (Paton *et al.*, 2001). But, these strains are thought to use other mechanisms for adhesion, like the LEE-negative O113:H12 STEC strain, shown to utilise the adhesin Saa (Paton *et al.*, 2001).

1.1.5.2 Attaching-and-effacing (A/E) lesions induce actin pedestal formation

A/E lesions induced by EHEC lead next to the loss of microvilli to localised actin polymerisation and to the formation of pseudopod-like “actin pedestals” beneath the attached bacteria (Figure 1.1-2). *E. coli* typically possess fimbriae, enabling the bacteria to adhere to host cells. The Type 1 fimbriae were described first and are the most common adhesins in *E. coli* (Duguid *et al.*, 1955; Welinder-Olsson and Kaijser, 2005) which mediate adherence to mannose-containing glycoproteins found on the surface of many eukaryotic cells (Johnson, 1991). However, the Stx producing EHEC strain O157 is unable to express type 1 fimbriae (Enami *et al.*, 1999), but adhere through their common pilus (ECP) (Croxen and Finlay, 2010) as well as through their type IV pilus, termed haemorrhagic coli pilus (HCP) (Figure 1.1-3, lower panel) (Xicohtencatl-Cortes *et al.*, 2007).

Attachment can also be enhanced by interactions of the bacterial outer membrane protein intimin with nucleolin, a surface-localised intimin receptor, whose expression is increased by Stx2 (see also section 1.1.4.3). The biggest impact on adherence, however, is mediated by the interaction of intimin to the effector protein Tir. Tir and a variety of other LEE and phage encoded effector proteins are translocated into the host cell via the T3SS (Figure 1.1-3, lower panel) (Hayward *et al.*, 2006). Therefore, a deletion of the *tir* gene results in a profound reduction in EHEC colonisation (~10.000-fold), as shown for the infant rabbit intestine (Ritchie and Waldor, 2005). Upon secretion, Tir is integrated into the plasma membrane in a hairpin-loop conformation and its extracellular domain serves as another receptor for intimin, thereby anchoring the bacterium to the host cell (Campellone and Leong, 2003). In this arrangement, both termini of Tir project into the host cell cytoplasm providing the recruitment sites for the host’s actin polymerisation machinery (Figure 1.1-3). The amino-terminal cytoplasmic domain is not strictly required for pedestal formation but modulates pedestal length (Campellone *et al.*, 2006). Its exact link to the actin cytoskeleton is still largely unclear. In contrast, the carboxy-terminal cytoplasmic domain (TirC) is essential for pedestal formation (Campellone *et al.*, 2006).

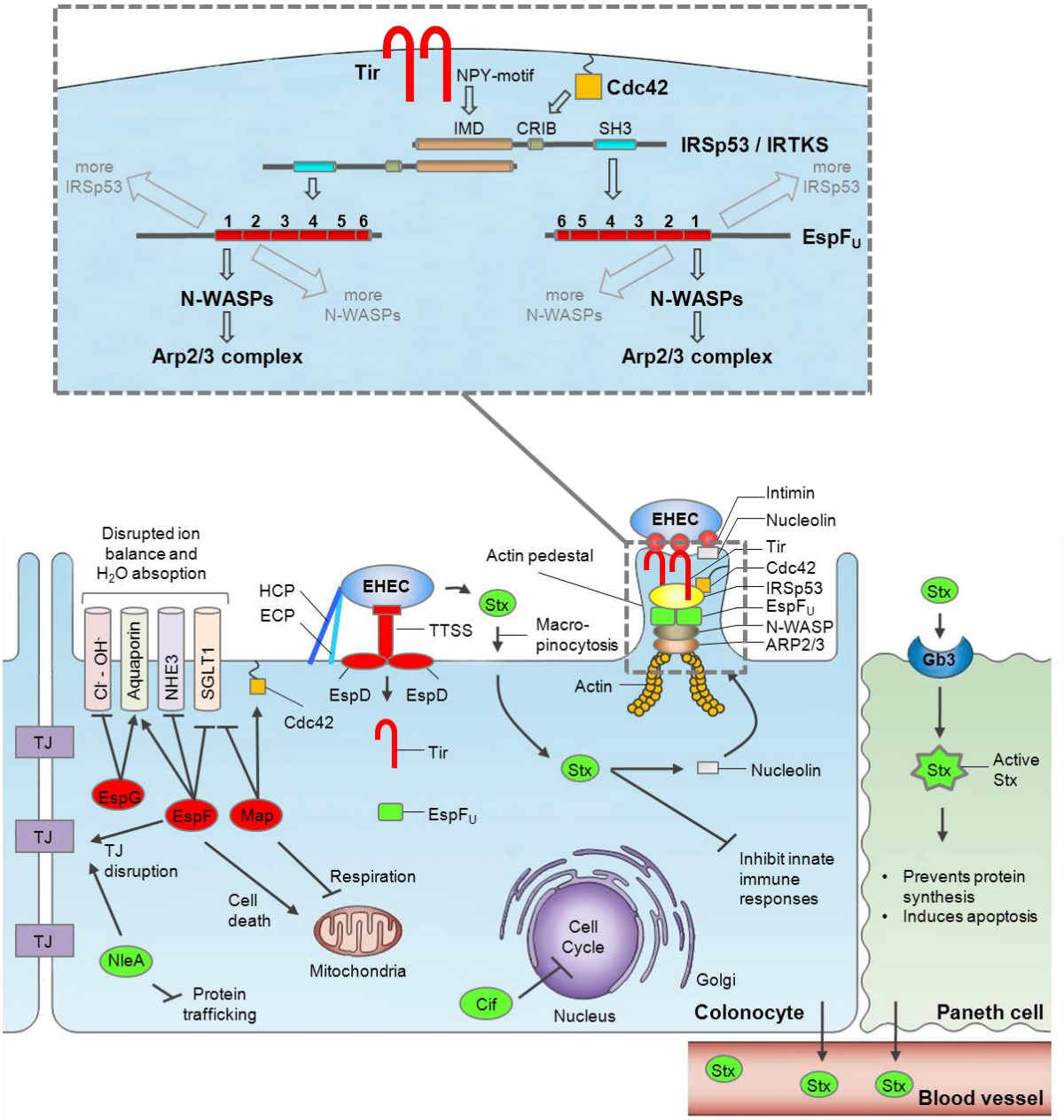

Figure 1.1-3: Schematic representation of A/E lesions during the course of EHEC infection according to Croxen *et al.* (2010) and Weiss *et al.* (2009), respectively. EHEC adhere to colonocytes in the large bowel through the *E. coli* common pilus (ECP) and the haemorrhagic coli pilus (HCP). The EHEC secrete a large arsenal of effector proteins into the host cell by a type III secretion system (T3SS). The effector protein Tir integrates into the plasma membrane and serves as a receptor for the EHEC cell surface protein intimin. The autoinhibited host protein IRSp53 (or its paralogue IRTKS) is activated by interacting with active Cdc42 and is thereby recruited to the membrane. Active IRSp53 interacts with Tir and EspF$_U$, another secreted effector protein. EspF$_U$ recruits and activates N-WASP, leading to ARP2/3 dependant actin polymerisation. Thereby, EHEC are linked to the actin cytoskeleton of the host cell, resulting in actin pedestal formation. Shiga toxin (Stx) released into the gut is internalized, possibly by macropinocytosis, into colonocytes, resulting in inhibition of the innate immune response and increased nucleolin expression, which enhances adhesion via additional intimin interactions. Stx is internalized into paneth cells via Gb3 receptor-mediated endocytosis. Upon enzymatic activation, the toxin prevents protein synthesis of the host cell, which can lead to apoptosis. After absorption, Stx can travel with the blood stream to other organs, causing severe damage especially to the kidneys. The effector proteins EspG, EspF and Map trigger a disrupted ion balance and inhibit water absorption, leading to watery diarrhoea. Inset: Schematic representation of the

domain interactions linking Tir_{EHEC} to Arp2/3 complex-mediated actin polymerisation during pedestal formation. The IMD and SH3 domains of IRSp53 or IRTKS interact with the NPY-motif of Tir_{EHEC} and the proline-rich repeats of $EspF_U$, respectively. The proline-rich repeats of $EspF_U$ are capable of interacting with N-WASPs, bound to the Arp2/3 complex which drives actin polymerisation, leading to a direct link of the extracellular pathogenic bacterium to the host cell's actin cytoskeleton. The LEE and prophage encoded effector proteins are shown in red and green, respectively.

$EspF_U$, as mentioned before, is also transported via the T3SS into the host cell. $EspF_U$ comprises a variable number of almost identical, proline-rich 47 amino acid repeats (Garmendia *et al.*, 2004; Sallee *et al.*, 2008) (Figure 1.1-3, inset), each including a hydrophobic helix that binds to and activates the autoinhibited N-WASP (Campellone *et al.*, 2004; Cheng *et al.*, 2008), leading to Arp2/3 dependant actin polymerisation. Furthermore, the proline-rich repeats are also capable of binding to the Src homology 3 (SH3) domain of either the host cell protein IRSp53 (insulin receptor substrate p53) (Weiss *et al.*, 2009) or its paralogue IRTKS (insulin receptor tyrosine kinase substrate) (Vingadassalom *et al.*, 2009). Due to multiple proline-rich repeats, $EspF_U$ was shown to simultaneously bind and activate multiple N-WASP proteins (Sallee *et al.*, 2008) and multiple SH3 domains. This allows the focal clustering of IRSp53 and Arp2/3 complexes and drives focal actin assembly (Weiss *et al.*, 2009), leading to pedestal formation.

IRSp53 has an SH3 domain, a central Cdc42/Rac interactive binding (CRIB) motif and an amino-terminal IRSp53-MIM homology domain (IMD) (see also section 1.1.6.1). The CRIB motif seems to be important for the activation of IRSp53, as IRSp53 exists in an autoinhibited state in the cytoplasm but becomes activated after the interaction of the CRIB domain with the membrane associated active form of Cdc42 (Ahmed *et al.*, 2009). The CRIB domain of IRSp53 is atypical, as it comprises less conserved amino acids compared to the typical CRIB of e.g. WASP (Krugmann *et al.*, 2001) and thus is also called "partial CRIB". Despite this, the CRIB of IRSp53 binds specifically to active Cdc42 but not to the activated Rac (Krugmann *et al.*, 2001; Scita *et al.*, 2008). The activation of the small Rho GTPase Cdc42 is mediated by the T3SS secreted effector protein Map (Figure 1.1-3) (Huang *et al.*, 2009). Map was shown to induce Cdc42 dependant cell surface filopodia (Kenny *et al.*, 2002) and to work as a guanine nucleotide exchange factor (GEF) that specifically activates Cdc42 (Huang *et al.*, 2009), recruiting the activated IRSp53 to the membrane (Ahmed *et al.*, 2009). Furthermore, Map acts as a multifunctional protein that also disrupts the mitochondrial structure and function (Ma *et al.*, 2006).

The IMD of the active IRSp53 interacts with TirC of EHEC, leading to a direct connection of the extracellular EHEC with the actin polymerisation machinery of the host cell (Vingadassalom *et al.*, 2009; Weiss *et al.*, 2009) and the recruitment of the N-WASP/Arp2/3 complex to the plasma membrane. The interaction surface within the IMD has been mapped

down to a stretch of nine amino acids with a central NPY motif (Campellone *et al.*, 2006). Substitutions with alanine demonstrated that each position of the NPY motif but none of the neighbouring positions are critical for actin pedestal formation. Interestingly, a conservative substitution of the tyrosine by phenylalanine attenuated pedestal formation but could not entirely abrogate it (Brady *et al.*, 2007). Despite the related structures of the Tir binding proteins IRSp53 and IRTKS, ectopic expression of IRSp53 mediates the formation of filopodia (Lim *et al.*, 2008), while cytoskeletal changes induced by IRTKS are more subtle (Millard *et al.*, 2005). In any case, these proteins link changes of plasma membrane morphology to actin cytoskeleton dynamics (Millard *et al.*, 2005; Scita *et al.*, 2008; Suetsugu *et al.*, 2006).

As mentioned above, along with Tir and $EspF_U$, a lot of other effector proteins are injected into the host cell, like Nle (also known as EspI) that reduces protein trafficking and also disrupts tight junctions, or the cycle inhibiting factor (Cif) that inhibits the cell cycle progression. Some effector proteins account for the primary clinical symptom of an EHEC infection, diarrhoea. Diarrhoea provides the EHEC with an advantage over competing flora of the intestine, since the intimately bound EHEC remain attached to the host's intestine and other less adherent bacteria are flushed away. The diarrhoea is mediated by the loss of absorptive surface due to the loss of microvilli, disrupted ion balance, and water absorption. This endothelial dysfunction is triggered by the T3SS translocated effector proteins like EspF and EspG, leading to a decreased activity of the Na^+/H^+ exchanger 3 (NHE3) (Hodges *et al.*, 2008) and the Cl^-/OH^- exchanger (Gill *et al.*, 2007), respectively (Figure 1.1-3, lower panel). The affected ion absorption leads to diarrhoea, as water moves according to the osmotic effect to areas of higher salt concentrations. Furthermore, both proteins induce mislocalisation of aquaporin water channels from the membrane into the cytoplasm (Guttman *et al.*, 2007). EspF, intimin, Map and Tir also cause the inhibition of the sodium-D-glucose co-transporter (SGLT1), a major water pump that accounts for approximately 4 to 6 litre water uptake from the intestine per day (Dean *et al.*, 2006; Meinild *et al.*, 1998). The inhibition of the intestinal serotonin transporter SERT (Esmaili *et al.*, 2009) and the upregulation of connexin 43 (Cx43) (Guttman *et al.*, 2011) were also shown to contribute to diarrhoea due to A/E bacterial infection (not shown in the figure). SERT is involved in the uptake of serotonin, a hormone regulating fluid and electrolyte secretion and absorption (Croxen and Finlay, 2010). Connexin 43 (Cx43) forms open connexon hemichannels in apical membranes of infected cells (Guttman *et al.*, 2011). However, up to now no T3SS effector proteins are identified to be associated in the SERT inhibition or Cx43 upregulation. But these dramatic effects of the secreted virulence factors on the epithelial cells, that also include the loss of tight junctions, lead unavoidably to severe diarrhoea. The initially watery

diarrhoea turns to bloody diarrhoea within the first weak (Figure 1.1-1), due to the mucosal infraction in response to the EHEC produced Shiga toxin (see section 1.1.4.4).

1.1.5.3 Comparison of EHEC and EPEC A/E lesions

Although EHEC and EPEC are closely related and have an almost identical LEE, EHEC injects around twice as many effectors into host cells as EPEC, even if most of these are redundant (Tobe *et al.*, 2006). These pathovars exhibit distinct adhesion mechanisms and pathways upon attachment that are exploited for inducing actin assembly, as shown for EHEC O157:H7 and EPEC O127:H6 or by other, more recently identified EHEC and EPEC strains (Frankel and Phillips, 2008). However, there are some correlations in particular between the prototypical strains EPEC 1 and the typical non-sorbitol fermenting EHEC O157, as shown in Figure 1.1-4 and discussed in the following. Both strains translocate Tir by the T3SS into the host. The Tir molecule is well conserved between EPEC, EHEC, and also other related bacteria such as *Citrobacter rodentium*. Only a short stretch of TirC around Tyrosine 474 is specific for EPEC-Tir. Phosphorylation of EPEC Y474 by host cell kinases (Phillips *et al.*, 2004; Swimm *et al.*, 2004) triggers binding of the SH2-domain of Nck proteins (Frese *et al.*, 2006; Gruenheid *et al.*, 2001), resulting in recruitment and activation of N-WASP and stimulation of Arp2/3 dependant actin polymerisation (reviewed in Hayward *et al.*, 2006). Thus, in contrast to EHEC, EPEC exploit Nck as part of their infection strategy (Frese *et al.*, 2006) and links the bacterium IRSp53 and $EspF_U$ independant to the hosts actin cytoskeleton.

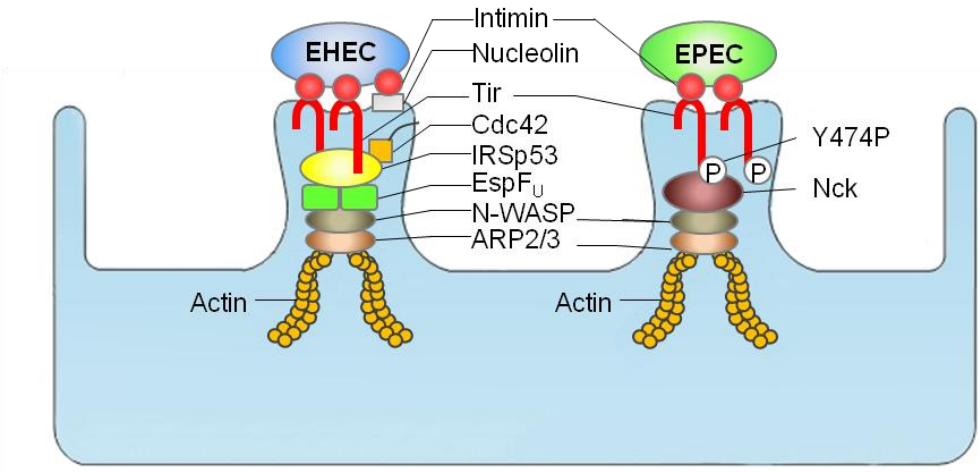

Figure 1.1-4: Comparison of the signalling cascade of the prototypical strains EHEC O157 (left) and EPEC 1 (right). While EHEC require IRSp53 and $EspF_U$ to link the bacterium to the host cell's actin cytoskeleton, EPEC exploit Nck as a linker to the actin cytoskeleton, which implicates the phosphorylation of Tir's Tyr474. Figure according to (Croxen and Finlay, 2010).

However, even though IRSp53 is present in EPEC but dispensable for EPEC pedestal formation, introduction of $EspF_U$ into EPEC leads to Nck-independant pedestal formation by using an $_{452}NPY_{454}$ motif of Tir similar to EHEC Tir (Brady *et al.*, 2007). Thus the canonical EPEC is capable of interacting with IRSp53, whereas no information are currently available about whether Tyr454 is phosphorylated or not when bound to IRSp53.

1.1.6 The domain organisation of IRSp53 and the homologue proteins IRTKS and FLJ22582

The proteins most closely related to IRSp53 are IRTKS and the poorly characterised FLJ22582. These proteins share the two main domains: IMD (see section 1.1.6.1) and SH3 (see section 1.1.6.2), but also a WASP homology 2 (WH2)-like domain (Figure 1.1-5), involved in mediating binding to G-actin (Scita *et al.*, 2008). The partial CRIB domain is a unique feature of IRSp53. FLJ22582 lacks also the partial CRIB but in addition the WW domain-binding motif (PPP*X*Y).

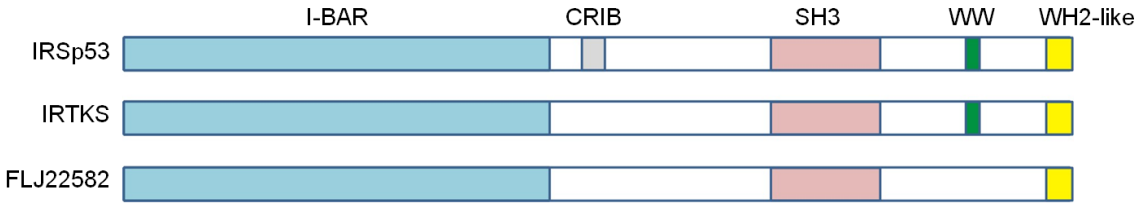

Figure 1.1-5: Domain organization of IRSp53 and the homologue proteins IRTKS and FLJ22582. The different domains of the proteins are indicated.

1.1.6.1 The IMD (I-BAR domain)

The alpha-helical IMDs are distant members of the Bin-Amphiphysin-Rvs (BAR) domain superfamily. Whereas canonical banana-shaped BAR domains bind membranes with their concave surface and induce invagination, IMDs are cigar-shaped and possess a convex surface that interacts with the cell membrane, leading to outward protrusions (Figure 1.1-6) like filopodia and are therefore referred to as inverse BAR or I-BAR domains (Frost *et al.*, 2009; Mattila *et al.*, 2007). Interestingly, IMDs also directly bind to and bundle F-actin (Disanza *et al.*, 2006; Millard *et al.*, 2005; Suetsugu *et al.*, 2006). However, two recent studies suggested that the isolated IMD itself does not crosslink actin filaments under physiological conditions (Lee *et al.*, 2007; Mattila *et al.*, 2007). The actin binding sites overlap, at least in part, with those that target lipid membranes (Millard *et al.*, 2005; Suetsugu

et al., 2006). IMDs deform PI(4,5)P_2-rich membranes (Mattila *et al.*, 2007; Saarikangas *et al.*, 2009) whereas the IMDs of IRSp53 or IRTKS interact mainly through electrostatic interactions with phospholipids of the membrane (Saarikangas *et al.*, 2009). In the resulting tubules, the IMDs remain dynamically associated with the inner leaflet and are thought to form an inner helical scaffold (Figure 1.1-6B) (Saarikangas *et al.*, 2009). It should be noted though, that the basic cluster of amino acids that confers lipid and potentially actin binding to IMDs (Millard *et al.*, 2005; Suetsugu *et al.*, 2006) is not required for formation of pedestals by EHEC bacteria (Weiss *et al.*, 2009).

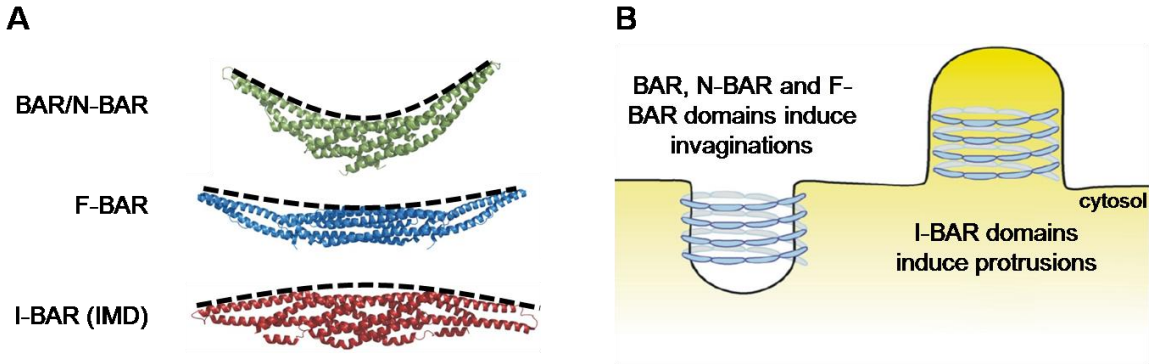

Figure 1.1-6: Hypothesis of BAR domain mediated membrane curvature. (A) Structural comparison of the three different types of BAR domains. Dashed lines indicate the assumed membrane binding site. (B) BAR domains lead according to their structure to the formation of invaginations (left) or protrusions (right). The BAR domains are hypothesized to exhibit a helical scaffold on the cytosolic side of the tubules. Figure (B): (Scita *et al.*, 2008).

1.1.6.2 The SH3 domain

The SRC homology 3 domain (SH3) plays a critical role in a wide variety of biological processes like increasing the local concentration of distinct proteins, altering the subcellular localisation of components of signalling pathways or mediating the assembly of large multiprotein complexes (Mayer, 2001). SH3 domains interact with proline-rich motifs of different proteins. The domain forms a typically β-barrel structure, with anti-parallel β-strands that are connected by variable loops. The small interaction surface of about 400Å² results in usually weak selectivity and affinities in the low micromolar range to the ligand (Kuriyan and Cowburn, 1997). However, affinity and specificity can be greatly enhanced by additional contacts between SH3 loop regions and residues of the ligand outside the proline-rich motif (Mayer and Saksela, 2004).

It is not surprisingly that SH3 domains exhibit also a high impact for the IRSp53 signalling, as they mediate the interactions to a huge variety of different ligands like WAVE2 (Miki *et al.*,

2000), Mena/VASP (Krugmann *et al.*, 2001), Eps8 (Disanza *et al.*, 2006), mDia (Fujiwara *et al.*, 2000), espin (Sekerkova *et al.*, 2003), DRLPA (Okamura-Oho *et al.*, 1999), Shank-1 (Bockmann *et al.*, 2002), and synaptopodin (Yanagida-Asanuma *et al.*, 2007). As the SH3 of IRSp53 and IRTKS also interacts with $EspF_U$ (Figure 1.1-3), it displays an essential component to physically link Tir and thus EHEC to the actin cytoskeleton of the host cell.

1.1.7 Therapeutic treatment of HUS

The treatment of Shiga toxin-producing bacteria with conventional anti-microbial therapies was often reported to be counterproductive since it may stimulate toxin production (Shimizu *et al.*, 2009; Wong *et al.*, 2000; Zhang *et al.*, 2000). Due to the worst ever outbreak of the very aggressive Stx2 producing HUSECO41 (O104:H4) in Germany 2011, there is a huge demand for therapeutic strategies to treat STEC infections and in particular the associated HUS. Currently no other therapeutic treatment other than fluid volume control, apheresis and dialysis are available to reduce or prevent renal failure. However, there are some promising approaches like the synthetic Shiga toxin binder “DAISY” (Mulvey *et al.*, 2003) or “SUPER TWIG” (Nishikawa *et al.*, 2002) which acts as a Gb3 analogue neutralizing the toxin. Probiotic bacteria with Gb3 mimics on their surface (Paton *et al.*, 2000) or monoclonal antibodies against Stx1 and Stx2 were also tested (Mukherjee *et al.*, 2002). Low purity of the commercially available Stxs are thought to be the reason for diverging results from related experiments, as contaminating substances like lipopolysaccharides are much more potent pro-inflammatory agents than Stxs (Obrig, 2010).

Treatment of HUSECO41 patients with the monoclonal antibody Eculizumab (Soaris; Alexion) to prevent the destruction of red blood cells, dramatically demonstrates the need for new therapeutic strategies. The antibody is developed to treat paroxysmal nocturnal haemoglobinuria, preventing integration of a pore forming terminal complement complex into erythrocytes which would cause haemolysis (Parker *et al.*, 2007; Rother *et al.*, 2007). But unfortunately so far only limited data is available which describes treatment of human HUS patients using this antibody (Gruppo and Rother, 2009; Lapeyraque *et al.*, 2011).

1.2 Aim of this work

IRSp53 and IRTKS were recently discovered to link the human pathogenic bacterium EHEC to the host cell actin polymerisation machinery by interactions of their IMD and SH3 domains to the translocated effector proteins Tir and $EspF_U$, respectively.

The aim of this study was to solve the crystal structure of the amino-terminal human IMD of IRSp53 in complex with a synthetic peptide derived from the EHEC effector Tir, shown to interact with IMD (Weiss *et al.*, 2009). The structure should uncover the unknown binding site on the surface of IMD for the NPY motif of the carboxy-terminus of Tir which was already proven to be essential for binding (Brady *et al.*, 2007).

Isothermal titration calorimetry should be used to give information about the thermodynamics and the stoichiometry of this interaction. Close inspection of the structure should facilitate design of IMD point mutants, predicted to be impaired in Tir binding, to verify the structural data. Furthermore, the mutants should identify the most essential residues of the IMD:Tir interaction and give information about their impact.

Hence, this study should reveal a detailed description of a new binding pocket on the surface of IRSp53's IMD, responsible for Tir recognition. The study should unveil how the NPY motif provides a central molecular scaffold, linking the bacterium and the host cells actin polymerisation machinery. Furthermore, the structure should reveal a novel binding site for NPY motifs, which is probably also implicated in interactions with other so far unknown interaction partners, effecting pathogenesis and/or actin cytoskeletal regulation.

1.3 Materials

All chemicals were purchased from the following companies: ApliChem, Gold Biotechnology, Fluka, Roth, Roche, Hampton Research, Merck, Millipore, Sigma-Aldrich and QIAGEN. The quality standard was *"pro analysis"* (p.a.).

1.3.1 Molecular weight standards

Precision Plus Protein Unstained Standards and Precision Plus Protein All Blue Standards (Biorad) were used for SDS-PAGE analysis.

1.3.2 Media, buffers and stock solutions

Table 1.3-1: Media, buffer and stock solutions used in this study.

Media	Composition
Luria Bertani (LB)	10 g/L Bactotryptone, 5 g/l NaCl, 5 g/L yeast extract
Super broth (SB)	900 ml of solution 1(12 g/L Bactotryptone, 24 g/L yeast extract, 0.4% (v/v) glycerol) were mixed with 100 ml of solution 2 (0.17 M KH_2PO_4 and 0.72M K_2HPO_4). The solutions were sterilized separately.
Buffers	**Composition**
8x SDS-sample-buffer	7% SDS (w/v), 20% (v/v) glycerol, 0.5 M β-mercapto ethanol, 0.25 mg/ml Coomassie Blue G-250, 10 mM Tris/HCL pH: 6.8
SDS-PAGE lower buffer (4x)	1.5 M Tris/HCl pH 8.8
SDS-PAGE upper buffer (4x)	0.5 M Tris/HCl pH 6.8, 0.4% (w/v) SDS
SDS-PAGE running buffer	25 mM Tris/HCl, 192 mM glycine, 0.1% (w/v) SDS
Protease inhibitors	Solution 1 (0.88 g PMSF, 10 mg Pepstatin A, 10 mg Chymostatin, dissolved in 10 ml DMSO) and solution 2 (1.57 g Benzamidine, 5 mg Leupeptin, 5 mg Bestatin, 5 mg Aprotinin, dissolved in 10 ml ddH2O) were sterilized by filtration, stored at -20°C and separately added to the protein solution in a ratio of 1:1000.

1.3.3 Bacterial strains

Table 1.3-2: Bacterial strains used in this study.

Bacterial strains	Genotype	Source
BL21(DE3)	F^- *ompT hsdS*$_B$($r_B^-m_B^-$) *gal dcm* (DE3)	Novagen
SCS1	*recA1 endA1 gyrA96 thi-1 hsdR17*($r_K^-m_K^+$) *supE44 relA1*	Stratagene
Rosetta2(DE3)	F^- *ompT hsdS*$_B$($r_B^-m_B^-$) *gal dcm* (DE3) pRARE2 (Cam)	Novagen

1.3.4 Plasmids for IMD-IRSp53 expression

pQTEV: A PCR fragment encoding amino acids 1-250 (MSLS…VASN) of human IRSp53 (UniProt BAIP2_HUMAN), cloned with BamH1 and NotI into a pQTEV (GenBank AY243506), was used for expression, resulting in IMD_{IRSp53} fused to an amino-terminal His_7-tag and TEV cleavage site with the sequence GSTMSLS…VASN upon cleavage.

pRARE: The pRARE plasmid (Novagene) encodes tRNA genes to improve the codon usage in *E. coli* of Arg, Ile, Gly, Leu and Pro, except for Arg CCA/CGG (Novy *et al.*, 2001). In this study, the pRARE plasmid was used to enhance protein expression of IMD-IRSp53 in the SCS1 cells.

1.3.5 IMD constructs

Table 1.3-3: IMD constructs used in this study.

Construct		plasmids	Bacterial strain
IMD_{IRSp53}wt	1-250	pQTEV/pRARE	SCS1
IMD_{IRSp53}L28E*	as above	as above	SCS1
IMD_{IRSp53}K108A*	as above	as above	SCS1
IMD_{IRSp53}R193S*	as above	as above	SCS1
IMD_{IRSp53}F196A*	as above	as above	SCS1
IMD_{IRSp53}K108A/R193S*	as above	as above	Rosetta2

*The IMD mutants used in this study were kindly provided by Prof. Dr. Theresia Stradal (Institute for Molecular Cell Biology, University of Münster, Germany).

1.3.6 Peptides

Table 1.3-4: Peptides used for ITC.

Peptides	Sequenze	SwissProt
EHEC-Tir-peptide	^{452}GTVQ**NPY**ADVKT^{463}	Q7DB77
EHEC-Tir-peptide-Y458F	^{452}GTVQ**NPF**ADVKT^{463}	as above
EPEC-Tir-peptide	^{448}SEVV**NPY**AEVGG^{459}	B7UM99
NPxY-peptide of the human insulin receptor	^{995}YS**NPEY**LSASDV^{1005}	P06213

1.3.7 Crystallisation screens

Initial crystallisation conditions were identified by using the following screens from QIAGEN: JCSG+, JCSG Core I, JCSG Core II, JCSG Core III, JCSG Core IV, The Classics, The PEGs, The PEGs II, The MPD, Cryos and The $AmSO_4$.

1.4 Methods

1.4.1 Production and purification of IRSp53-IMD and IRSp53-IMD mutants

E. coli SCS1 or Rosetta2 (Table 1.3-3) cells bearing the IMD-IRSp53 expression plasmid and pRARE were grown in SB medium supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol at 37°C to an OD of 1.5. Protein expression was induced with 1 mM IPTG at 37°C for 4 hr. Harvested cells were disrupted with a high pressure cell disrupter (Constant Systems) in 20 mM Hepes, pH 7.5, 100 mM NaCl, 10 mM imidazole, 1 mM TCEP and protease inhibitors. After centrifugation (35,000×g, 45 min), the soluble protein was purified by nickel affinity chromatography (5 ml HiTrap Chelating HP column, GE Healthcare) and eluted with an imidazole gradient (10-500 mM) over 20 column volumes. The protein was cleavage with his-tagged TEV protease. Uncleaved material and the protease were removed by nickel rechromatography, followed by gel filtration (HiLoad 16/60 Superdex 75 prep grade column, GE Healthcare) in 20 mM HEPES-NaOH, 100 mM NaCl, 0.1 mM EDTA and 1 mM DTT, pH 7.5. The protein purification for isothermal titration calorimetry (section 1.4.3.5) was carried out with exception of the His_7-tag truncation.

1.4.2 IRSp53-IMD:Tir complex formation

A peptide with the sequence ^{452}GTVQ**NPY**ADVKT^{463}, derived from EHEC Tir (see section 1.3.6), was synthesized and HPLC-purified in-house by Dr. Werner Tegge. The peptide contains a carboxy-terminal amide and was used for complex formation and crystallisation (see section 1.4.4). The peptide buffer was exchanged against 20 mM HEPES-NaOH, 100 mM NaCl, 0.1 mM EDTA and 1 mM DTT, pH 7.5. The complex was prepared by adding a 2-fold molar excess of peptide to purified IMD with a final concentration of 11 mg/ml.

1.4.3 Analytical methods

1.4.3.1 Photometric quantification of protein concentrations

The protein concentrations were determined at 280 nm by using a NanoDrop ND-1000 photometer (peqlab) according to the Lambert-Beer law (Equation 1.4-1), with the extinction (E), the intensity of the transmitted beam (I_1), the intensity of the incident beam (I_0), the extinction coefficient (ε_λ), the protein concentration (c), and the path length (d). The extinction coefficients were calculated using the software ProtParam (www.ExPASy.org).

$$E_\lambda = -lg\left(\frac{I_1}{I_0}\right) = \in_\lambda \cdot c \cdot d$$

Equation 1.4-1

1.4.3.2 Discontinuous SDS-PAGE

SDS polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyse protein expression and quality (Laemmli, 1970). Protein samples were denatured in 8x SDS sample buffer for 5 min at 95 °C and applied next to a molecular standard to an SDS-gel (Table 1.4-1). The electrophoresis was initially performed at 120 V until the samples were concentrated to a thin line in the stacking gel. The electrophoretic separation in the running gel was performed at 160 V. The gels were briefly washed in dH_2O and subsequently stained with InstantBlue (Biozol).

Table 1.4-1: SDS gel composition. Specifications according six gels.

Solutions	Running gel 12%	Stacking gel 5%
Acrylamide/bisacrylamide 30 % (w/v)	12 ml	1.5 ml
4x lower buffer	7.6 ml	
4x upper buffer		2.5 ml
10% SDS	0.3 ml	
H_2O	10 ml	5.9 ml
TEMED	40 µl	30 µl
25% APS	60 µl	30 µl

1.4.3.3 Dynamic light scattering (DLS)

Dynamic light scattering (DLS) is a tool to determine size distribution profiles of particles or proteins in solution, providing information about the monodispersity or polydispersity and thus about complex formation, protein aggregation or degradation. The size distribution profile gives also allusion to the crystallisability, whereas proteins with a monodispersity of 70-80% were shown to have a high probability of producing some kind of crystals (Bergfors, 1999). DLS determines the scattering intensities of particles exposed to a focused laser beam as a function of time. The resulting fluctuation of the intensities, due to the Brownian Motion, depend on the particle size. Increasing particle size leads to decreased Brownian Motion and increased half-width of the peak maxima. The diffusion constant can be calculated by applying an autocorrelation function, leading to the possibility to calculate the hydrodynamic radius of the particles, giving information about the size distribution profile of particles in the solution.

For DLS the protein samples with a concentration of 0.5 – 1 mg/ml (leading to approximately 10% scattering of the laser beam) were filtrated with a 0.2 µm filter to avoid the interference of dust. DLS measurements were carried out at room temperature by using a DynaPro Titan (Wyatt Technologies) and a laser power of at least 1,000,000 counts/sec. Protein solutions with a polydispersity <30% were used for crystallisation.

1.4.3.4 Mass spectroscopy (MS)

Protein samples from SDS-PAGE were analysed in-house by Dr. Manfred Nimtz and Undine Felgenträger using MALDI-TOF-MS.

1.4.3.5 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is a technique to study biomolecular interactions by the direct measurement of protein binding energetics (Leavitt and Freire, 2001). By recording the heat of reaction (binding enthalpy, ΔH), directly proportional to complex formation, characteristic parameters can be calculated. Numerical approximation to a curve of cumulated ΔH (ΔH_{cum}) versus the total ligand concentration (B_t) according to Equation 1.4-2 facilitates the determination of the enthalpy of formation (ΔH_f) and the association constant (K_a). The dissociation constant (K_d) corresponds to the reciprocal association constant according to Equation 1.4-3. The Gibbs energy changes (ΔG) and entropy changes (ΔS) are calculated according to the Van't Hoff and Gibbs-Helmholtz equation (Equation 1.4-4 and Equation 1.4-5).

$$\Delta H_{cum} = \frac{V \cdot \Delta H_f}{(2 \cdot K_a)} \cdot [K_a \cdot (C + Bt) + 1 - ({K_a}^2 \cdot [C - Bt]^2 + 2 \cdot K_a \cdot [C + Bt] + 1)^{0.5}] \quad \textbf{Equation 1.4-2}$$

$$K_d = \frac{1}{K_a} \quad \textbf{Equation 1.4-3}$$

$$\Delta G = -R \cdot T \cdot ln(K_a) \quad \textbf{Equation 1.4-4}$$

$$\Delta G = \Delta H - T \cdot \Delta S \quad \textbf{Equation 1.4-5}$$

To determine these binding parameters by ITC, a MCS-ITC calorimeter (Microcal, Northhampton, MA, USA) was used. Peptides and IMD were dialyzed against 20 mM HEPES-NaOH, 100 mM NaCl, 0.1 mM EDTA and 0.8 mM tris(2-carboxyethyl)phosphine (TCEP), pH 7.5. During titration, 2 mM peptide was injected into 0.1 mM of His-tagged IMD protomers at 24°C in 10 µl steps up to a threefold molar excess. The mixing heat of the peptide, measured by injecting peptide into buffer, was subtracted from the signals obtained from the binding reactions. ITC data of three independant measurements were analysed using Origin V7.0 with Microcal ITC add-on. Peptides ^{452}GTVQ**NPY**ADVKT^{463}, derived from EHEC Tir, GTVQ**NPF**ADVKT, ^{448}SEVV**NPY**AEVGG^{459}, derived from EPEC Tir (SwissProt B7UM99), and a sequence of the human insulin receptor ^{995}YS**NPEY**LSASDV^{1005} (SwissProt P06213) were synthesized in-house with a caboxy-terminal amide and HPLC-purified by Dr. Werner Tegge.

1.4.4 Crystallisation by vapour diffusion

The vapour diffusion technique is the most commonly used method for crystallisation. This method based on the presence of a reservoir with precipitant containing mother liquor, absorbing water from a crystallisation drop in a closed system. Water vapour diffuses from the crystallisation drop, a mixture of protein solution and mother liquor, into the higher concentrated reservoir, until equilibrium is reached. The diffusion leads to a gradual supersaturation in the crystallisation drop and thus facilitates entropic driven crystallisation (Rupp, 2010).

Prior crystallisation, the concentrated an purified protein of the IRSp53-IMD:Tir complex was centrifuged (20 min., 18.000g, 4°C) to separate possible precipitates from the protein solution. Crystallisation was performed at 19°C using the vapour phase diffusion method, either on a sitting drop for initial screening or on a hanging drop for optimisation.

1.4.4.1 Sitting-drop vapour diffusion

The sitting-drop vapour phase diffusion method was used for initial screening in 96-well format, facilitating the use of robotic systems to mix small amounts of pure protein with screen conditions (section 1.3.7).

200 nl of the purified IMD-IRSp53:Tir protein was mixed with an equal amount of the screening conditions in a droplet by using a pipetting robot (Mosquito robot, TTP or HoneyBee 961, DIGILAB Genomic Solutions Inc.). The crystallisation was carried out in 3 x 96 well plates (Greiner), whereas two wells were loaded with different protein concentrations (11 mg/ml and 5.5 mg/ml) and the third was loaded with a buffer/precipitant control, to check for salt crystal formation. The 96-well plates were sealed using a Manco^{TM} Crystal Clear tape (Jena Bioscience) and incubated at 19°C.

1.4.4.2 Hanging-drop vapour diffusion

The hanging drop vapour phase diffusion method was used for optimisation in a 24-well format, by varying the pH, the protein, or precipitant concentration. Protein solution was mixed on a cover slip with an equal amount of mother liquor and placed on a well with a greased rim upside down over 500 µl of mother liquor.

The IMD-IRSp53:Tir complex was crystallised by mixing 2-3 µl of concentrated protein (11mg/ml) and 2-3 µl of mother liquor (16% (w/v) PEG 3350, 0.3 M $(NH_4)_2SO_4$), incubated at 19°C.

1.4.4.3 Microseeding

The Seeding technique was used to prevent microcrystal showers and to optimise crystal growth and quality. Microcrystal showers are caused or at least favoured by a crystallisation setup in a region of high supersaturation and spontaneous nucleation or by an experiment that proceeds far into the nucleation zone. This kind of setup often results in the formation of many microcrystals, consuming the protein in the crystallisation drop and preventing the formation of only a few big crystals. Lowering the concentration may prevent these microcrystals, but increase the time for nucleation. Thus it is sometimes worthwhile to identify the amount of seeds and the crystallisation conditions, often slightly reduced conditions, to optimise crystal quality by microseeding. The strategy of microseeding is to place micro seeds, e.g. from a crashed protein crystal, into a low supersaturated protein-precipitant solution, where spontaneous nucleation is improbable but nucleation growth of the seeds will occur, leading in theory to only a few big crystals.

5-10 IMD-IRSp53:Tir crystals, resulting from the hanging-drop method, were crushed in 100 µl of mother liquor (16% (w/v) PEG 3350, 0.3 M $(NH_4)_2SO_4$) and were diluted ten times 1:10. IMD-IRSp53:Tir solution was mixed with an equal amount of each dilution and incubated according to the hanging-drop vapour diffusion method.

1.4.5 Data collection, structure determination and refinement

To prevent or minimise crystal damage by radiation and thermal vibrations, the crystals were transferred into reservoir solution with additional 20% (v/v) PEG 400 before flash freezing in liquid nitrogen and data collection at 100 K. The data sets were collected at beamline 14.2 of the synchrotron storage ring BESSY II in Berlin. Images were indexed and processed with the XDS program package (Kabsch, 1993) and the structure was solved by molecular replacement using CCP4 MOLREP (Vagin and Teplyakov, 1997) with PDB entry 1Y2O as the search model. REFMAC5 (Murshudov *et al.*, 1997) was used for refinement. TLS (Translation/libration/screw) (Painter and Merritt, 2006) parameters were included to allow domain movements to be refined and Coot (Emsley and Cowtan, 2004) was used for manual model building. The validation of the structure was performed using MolProbity (Chen *et al.*, 2010; Davis *et al.*, 2007).

1.4.6 Bioinformatics

1.4.6.1 Sequence analysis

Protein sequences were aligned with EMBOSS emma/ClustalW2 (Larkin *et al.*, 2007) and displayed with ESPript 2.2 (Gouet *et al.*, 2003). Disordered protein regions were predicted by DISOPRED2 (Ward *et al.*, 2004). The proportion of human protein sequences in SwissProt containing the NPY motif was determined with Scansite (Obenauer *et al.*, 2003).

1.4.6.2 Structural analysis

Schematic diagrams of atomic interactions were calculated with LIGPLOT (Wallace *et al.*, 1995). Buried surface areas were calculated with PISA (Krissinel and Henrick, 2007). Root mean square deviations (r.m.s.d.) between common C_α-positions were calculated with ProFit using the McLachlan algorithm (McLachlan, 1982).

1.4.6.3 Figure preparation

Figures were prepared with PyMOL (http://pymol.org), LIGPLOT (Wallace *et al.*, 1995) and CorelDRAW X (Corel, Ottawa, Canada).

1.5 Results

1.5.1 Production and purification of IMD-IRSp53

The IMD of IRSp53 was expressed and purified according to section 1.4.1, using immobilised metal ion affinity chromatography (IMAC) for initial purification. Elution of the His_7-tag bound IMD by a gradient of imidazole, leads already to a high purity of the 31.7 kDa protein due to a late elution of the IMD at approximately 450 mM imidazole (Figure 1.5-1). The yield of IMAC purified IMD was approximately 40 to 80 mg per litre culture.

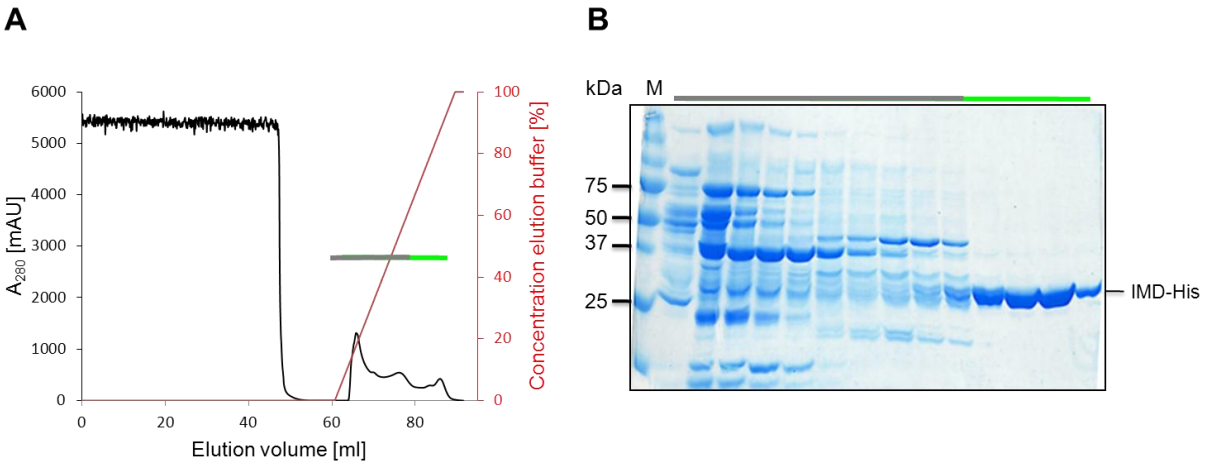

Figure 1.5-1: IMAC purification of IMD-IRSp53. Bacterial lysate containing the IMD-IRSp53 was loaded on a 5 ml HisTrap HP column and eluted with a gradient of 10 to 500 mM imidazole (A). The grey and green line depicts the fractions analysed by SDS-PAGE and Coomassie Blue staining (B). The fractions marked green were pooled for further purification. M: protein standard.

The His_7-tag was truncated from IMD using His_6-tagged TEV protease. The truncation avoids steric hindrance of crystallisation by the tag and minimises disordered regions, reducing unfavourable loss of entropy of the IMD during crystallisation. Nickel rechromatography was used to separate the His_7-tag and protease from the IMD after cleavage, and gel permeation chromatography was used for buffer exchange and final purification. Due to the high purity of the protein after the first purification step, the gel filtration column was loaded with the maximum of the recommended sample volume of 3% column volume, resulting in an expanded protein peak, followed by a second peak which can be attribute to nucleic acids (Figure 1.5-2A). SDS-PAGE was used to verify the purity and removal of the ~3 kDa His_7-tag (Figure 1.5-2B). Prior crystallisation trials, the pooled protein fractions of the gel filtration were analysed by MALDI-TOF-MS (section 1.4.3.4) and by DLS (section 1.4.3.3), to ensure the identity of the 28.8 kDa IMD and its monodispersity, respectively.

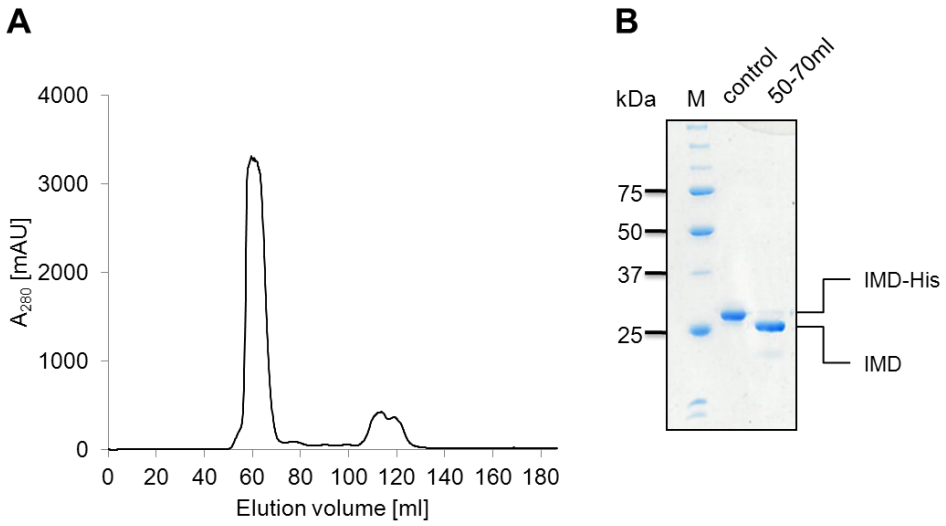

Figure 1.5-2: Gel filtration. The IMD elutes at about 62 ml from a Superdex 75 HR 16/60 column (A). The peak fraction (50-70 ml) was analysed by SDS-PAGE and Coomassie Blue staining (B), with His_7-tagged IMD protein of the IMAC as a control to verify the successful truncation of the approximately 3 kDa His_7-tag. M: protein standard.

1.5.2 Production and purification of IMD-IRSp53 mutants

The IMD-IRSp53 mutants used in this study (Table 1.3-3) were expressed and purified according to section 1.4.1 with yields and purification profiles during the IMAC and gel filtration similar to the wild type IMD (section 1.5.1). SDS-PAGE of purified IMD mutants is shown in Figure 1.5-3.

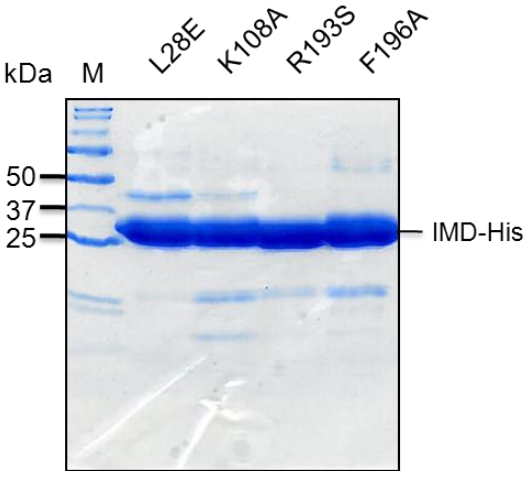

Figure 1.5-3: SDS-PAGE of the gel filtration purified His_7-tagged IMD mutants. The different mutants are indicated. M: protein standard.

1.5.3 Crystallisation of the IMD-IRSp53:Tir complex

Crystallisation trials were performed with either unbound IMD or IMD in complex with an EHEC or EPEC Tir derived synthetic dodecapeptide. The complexes were formed according to section 1.4.2. Crystallisation conditions were initially screened in 96-well format by using the sitting-drop vapour phase diffusion method (section 1.4.4.1). Using the JCSG screens, multiple crystallisation conditions for IMD, working as a control, and the IMD in complex with either an EHEC or EPEC peptide were identified (Table 1.5-1). However, only the IMD in complex with the EHEC derived Tir dodecapeptide could be optimised to well diffracting crystals. Even though, crystals appear already after one day using the optimised condition of 0.3 M ammonium sulphate and 16% PEG 3350 as mother liquor, and microseeding to optimise crystal growth, well diffracting crystals needed to grow up to one month at 19°C to reach a proper size (Figure 1.5-4) and quality for diffraction up to 2.1 Å (measured at the beamline 14.2 of the synchrotron storage ring BESSY II in Berlin).

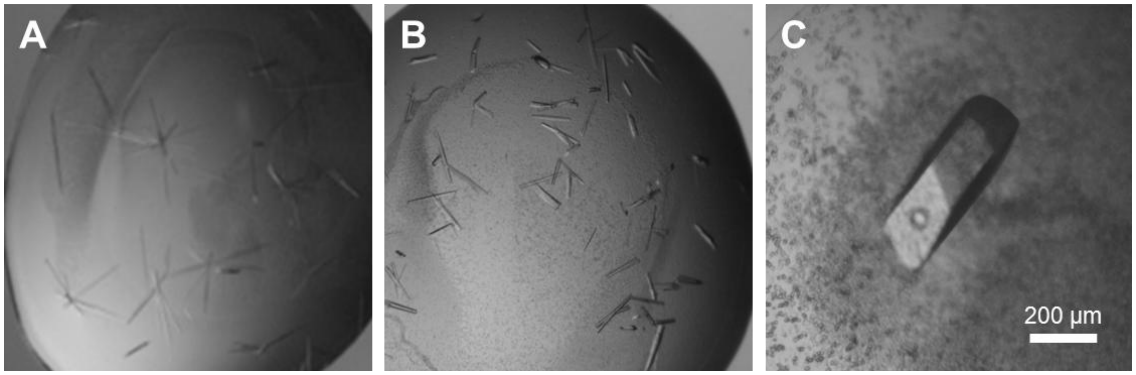

Figure 1.5-4: Optimisation of the crystallisation condition of IMD in complex with the EHEC derived Tir dodecapeptide. Initial optimisation of the crystallisation conditions by the hanging-drop vapour diffusion technique leads to conjoined needles (A). Further optimisation by using the microseeding technique leads to single occurring needles (B). Well diffracting crystal need to grow up to one month at the optimised condition and a temperature of 19°C.

1.5.4 Crystal structure of the IMD-IRSp53:Tir complex

The crystal structure of the complex of the IMD of IRSp53 with a dodecapeptide derived from EHEC Tir (GTVQ**NPY**ADVKT, residues 452-463) was solved by molecular replacement and refined at a resolution of 2.1 Å. The crystallographic data are summarized in Table 1.5-2. IMD forms a 170 Å long, cigar-shaped homodimer formed by a bundle of parallel coiled coils of kinked α-helices (Figure 1.5-5). The two protomers of the dimer measure about 120 Å in length and 30 Å in diameter and bury a surface area of 2,600 $Å^2$. The IMD structure is in accord with previously published IMD structures (Lee *et al.*, 2007; Millard *et al.*, 2005;

Suetsugu *et al.*, 2006), with a calculated r.m.s. deviation of 0.95 Å to Protein Data Bank (PDB) entry 1Y2O (218 common C_α-positions aligned).

Table 1.5-1: Crystallisation conditions for IMD, IMD in complex with an EHEC and EPEC Tir derived peptide. Optimised crystallisation condition used for IMD-Tir_{EHEC} crystallisation is highlighted in blue.

	Crystallisation conditions	Crystal form
IMD	• 20% (w/v) PEG 3350, 0.2 M lithium nitrate	needles
	• 20% (w/v) PEG 3350, 0.2 M calcium acetate	spheres
IMD+Tir_{EHEC}	• 0.1 M hepes pH 7.5, 40% (v/v) PEG 400, 0.2 M calcium acetate	spheres
	• 0.1 M sodium cacodylate pH 6.5, 40% (v/v) PEG 600, 0.2 M calcium acetate	spheres
	• 0.1 M MES, 1.6 M ammonium sulphate, final pH 6.0	spheres
	• 0.1M citric acid, 30% (w/v) PEG 6000, final pH 4.0	spheres
	• 0.2 M ammonium sulphate, 20% (w/v) PEG 3350	needles
IMD+Tir_{EPEC}	• 0.1 M sodium cacodylate pH 6.5, 40% (v/v) MPD, 5% (w/v) PEG 8000	spheres
	• 0.1 M sodium phosphate-citrate pH 4.2, 40% (w/v) Ethanol, 5% (w/v) PEG 1000	needles
	• 0.2 M calcium acetate, 0.1 M imidazole pH 8.0, 20% PEG 1000	spheres
	• 0.2 M lithium sulphate, 0.1 M sodium acetate pH 4.5, 50% PEG 400	spheres

1.5.4.1 Both Tir molecules bound by the IMD point into the same direction

The electron density map clearly reveals the position of the bound Tir-derived peptide (Figure 1.5-6A). The IMD homodimer binds two Tir-derived peptides, one at the centre of each protomer (Figure 1.5-5). The peptides point into the same direction, almost parallel to the two-fold rotation axis of the dimer and perpendicular to the IMD cigar shape. The carboxy-termini of the peptides point to the site of the IMD that is assumed to face the cytoplasm in the membrane bound state. Peptides are located far apart with 91 Å and 61 Å distances between their carboxy- and amino-termini, respectively. Each peptide buries 564 $Å^2$ of the IMD surface, which is close to the average of buried surface areas in peptide-protein interactions (London *et al.*, 2010). The whole length of the Tir-derived peptide is in close proximity to the domain, with the two Tir valine side-chains (V454, V461) oriented inwards, thus maximising solvent entropy upon binding. All amino acid residues of the peptide and the domain are well defined in the electron density map, except for the loop between helices H2 and H3 and residues 234-250 encoding helix H4.

Table 1.5-2: Summary of crystallographic analysis.

Data collection	
Space group	$C222_1$
Cell dimensions a, b, c (Å)	88.35, 187.36, 36.81
X-ray source	Beamline 14.2, BESSY II, Berlin
Wavelength (Å)	0.9184
Resolution range (Å)	9.44-2.11
Last shell (Å)	2.16-2.11
R_{merge} *(%)*	7.7 (53.2)
Observations	231684 (9729)
Unique reflections	18190 (1343)
Mean (I)/s.d. (I)	20.13 (3.13)
Completeness	99.8 (100)
Multiplicity	12.7 (7.2)
Structure refinement	
Resolution range (Å)	9.44-2.11 (2.16-2.11)
R_{work} (%)	23.3
R_{free} (%)	27.6
Total number of	
Non-hydrogen atoms	1991
Protein atoms	1837
Peptide atoms	103
Water molecules	46
Sulfate molecules	1
r.m.s.d.	
Bond length (Å)	0.022
Bond angle (deg)	1.90
Main-chain B-factors ($Å^2$)	39.4
Side-chain B-factors ($Å^2$)	43.5
Wilson B-factor ($Å^2$)	41.8
Average B-factor protein atoms ($Å^2$)	41.5
Average B-factor peptide atoms ($Å^2$)	27.0
Average B-factor solvent atoms ($Å^2$)	40.8
Ramachandran statistics	
Most favoured regions (%)	99.1
Allowed regions (%)	0.9
Disallowed regions (%)	0

Values in parentheses are for the highest resolution shell.

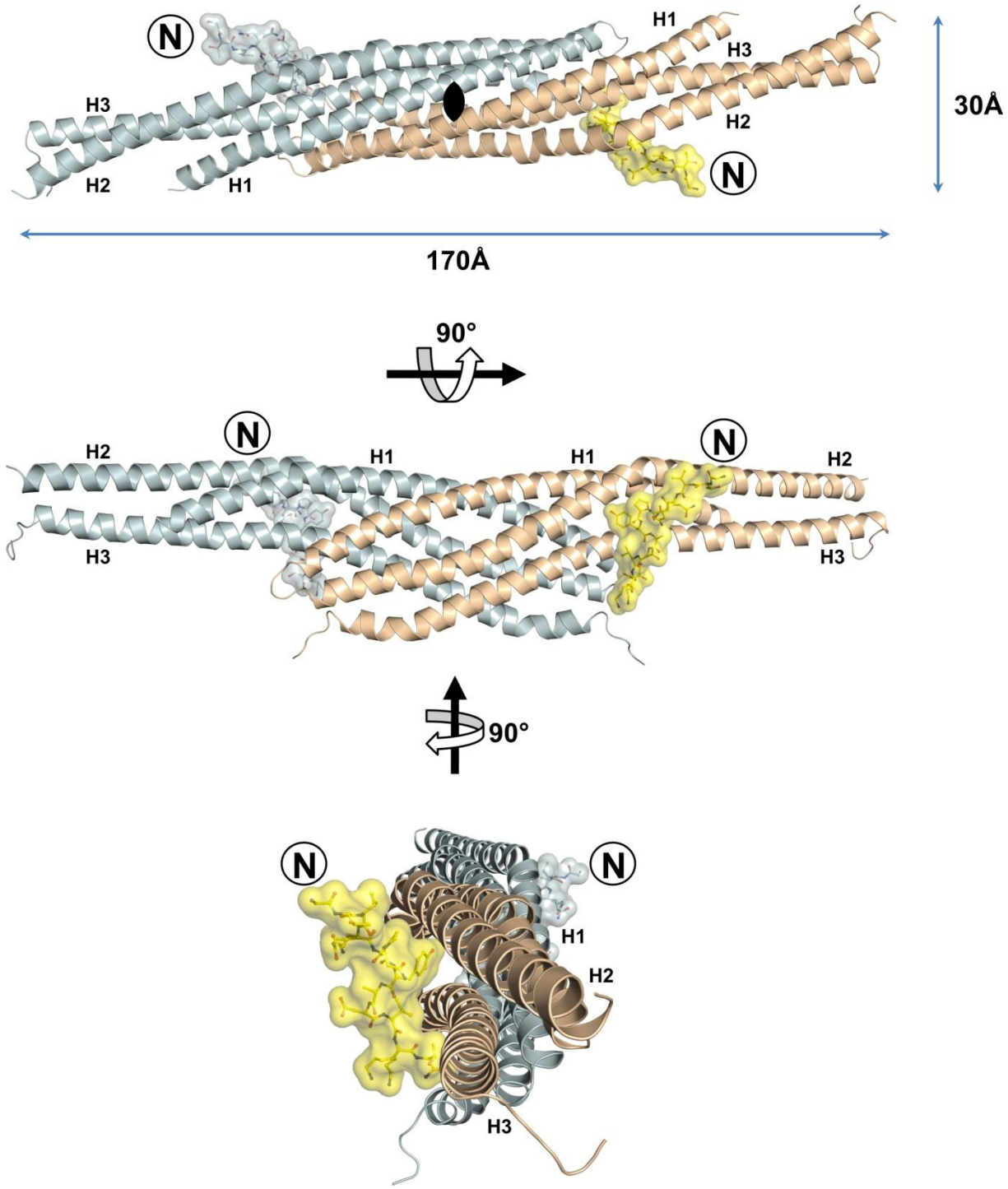

Figure 1.5-5: Complex of IMD_{IRSp53} with the EHEC Tir-derived peptide. IMD_{IRSp53} binds a 12-residue peptide derived from the EHEC virulence factor Tir, incorporating the NPY motif. Ribbon diagrams of the two subunits of the IMD homodimer are drawn in cyan and wheat. Two identical peptide molecules are coloured blue and yellow, respectively. Peptide amino-termini are labelled with an encircled N.

1.5.4.2 A novel peptide binding site of IRSp53 recognises the NPY motif

In the structure of the IMD:Tir complex, the NPY motif of Tir interacts with a large hydrophobic binding pocket in the central region of the IMD_{IRSp53} protomer (Figure 1.5-6A). This pocket is also present in the free IMD structure reported by Suetsugu *et al.* (PDB entry 1WDZ). On the contrary, in the structure reported by Millard *et al.* (PDB entry 1Y2O), the binding pocket of one of the two protomers is partly occluded by the side-chain of Lys108 (Figure 1.5-6C,D). The reason for this difference remains unclear.

All three side-chains of the NPY motif (456-458) are specifically bound by IMD_{IRSp53}. The conformation of the peptide backbone is defined by two overlapping type I β-turns (456-459, 458-461), which are stabilised by three intramolecular hydrogen bonds and an ordered water molecule (Figure 1.5-6B, Figure 1.5-7). The pyrrolidine ring of Pro457 of the NPY motif further rigidifies the backbone and is accommodated in a shallow hydrophobic groove on the IMD surface (Figure 1.5-6A,D).

The side-chain of Tyr458 of the NPY motif is accommodated in a deep hydrophobic pocket and is sandwiched between the hydrophobic portions of the IMD Arg193 and Lys108 side-chains (Figure 1.5-6A,B). Two ordered water molecules are bound in the interior of the pocket and link the Tyr458 hydroxyl group to helices H1, H2, and H3 of IMD (Figure 1.5-7, Figure 1.5-6B). Substitution of Tyr458 by phenylalanine would compromise these water-mediated interactions, which rationalises the reduced pedestal formation of the EHEC-Tir Y458F mutant (Brady *et al.*, 2007).

Recognition of Asn456 of the NPY motif is based on the unique hydrogen bonding potential of its side-chain (Figure 1.5-7). The side-chain's carbonyl oxygen is hydrogen bonded to a nitrogen of the guanidinium group of IMD Arg193 and accepts an additional, intrapeptide hydrogen bond from the backbone nitrogen of Tir Tyr458, an i to i+2 interaction that often stabilises Asn-Pro-containing β-turns (Wilson and Finlay, 1997). The side-chain amide nitrogen of Asn456 is hydrogen bonded to IMD Glu189. Hydrophobic contacts of the Asn456 side-chain to Tyr115 further strengthen the interaction. Clearly, asparagine is essential at this position, as no other residue could fully satisfy all of the observed bonding constraints.

Arg193 of IMD_{IRSp53}, part of the signature sequence ^{189}EERR**R**^{193} of the IMD family (Yamagishi *et al.*, 2004), appears to be the most critical residue for the recognition of the NPY motif. In the free domain, its side-chain is already oriented in the correct conformation for binding to Tir Asn456 and Tyr458 by a salt bridge between its guanidinium group and Asp112. This is in contrast to Lys108, which has a variable side-chain conformation in the absence of the peptide.

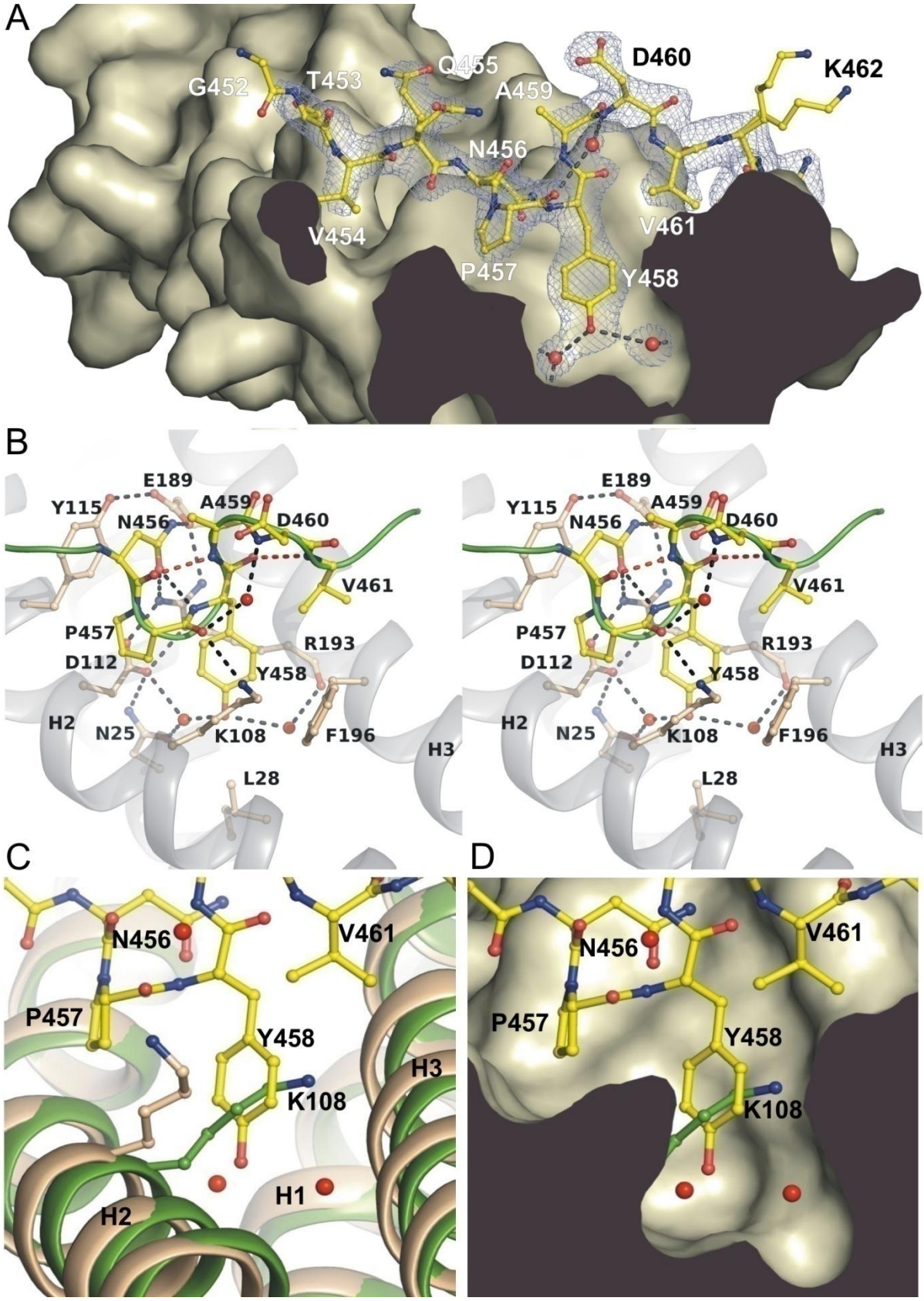

Figure 1.5-6: **Atomic details of the IMD_{IRSp53}:Tir complex.** (A) Profile of the binding pocket. The Tir-derived peptide ^{452}GTVQ**NPY**ADVKT^{463} is drawn in yellow with its electron density map in blue. Dashed lines indicate hydrogen bonds. Tyr458 of the NPY motif is centred by two water molecules (red) in the interior of the binding pocket. (B) Stereo view of the interactions of the Tir-derived peptide comprising the NPY motif (456-458). Residues involved in the complex interface and in β-turn

formation are labelled. The green line represents the backbone of the peptide. Hydrogen bonds involved in β-turn formation are shown as red dashed lines. (C and D) Superposition of the free IMD structure of Millard *et al.* (green, PDB 1Y2O) and the peptide-bound IMD (wheat). The side-chain of Lys108 blocks the binding pocket in the free structure. The conformation of the side-chain of Lys108 in the free IMD structure of Suetsugu *et al.* is similar to the complex (1WDZ, not shown).

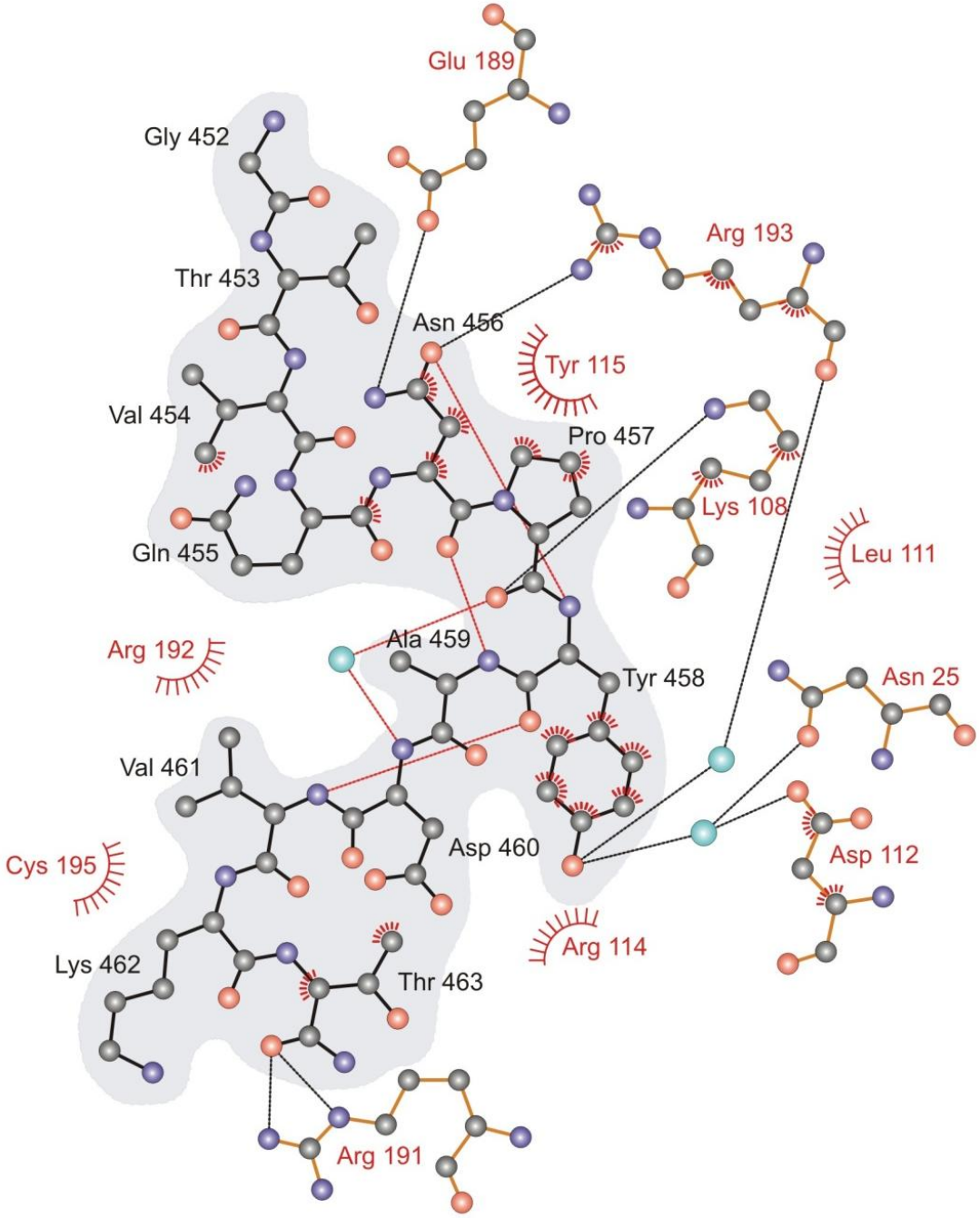

Figure 1.5-7: Schematic diagram of atomic interactions. EHEC Tir and the IMD's binding pocket are drawn with black and red labels and bonds, respectively. Cyan spheres represent water and dashed lines hydrogen bonds. Hydrophobic interactions are indicated by spoked arcs and atoms with spokes. An ordered water molecule and three intra-molecular hydrogen bonds (red) stabilise the peptide's two overlapping β-turns. Oxygen atoms are coloured red, nitrogen atoms blue and carbon atoms black.

In the structure of the free IMD homodimer of Millard *et al.* (PDB 1Y2O), Arg191 binds to Asp232 of the loop linking the opposing protomer's helices H3 and H4 (Figure 1.5-8A). In complex to Tir, the Arg191 guanidinium group contacts the oxygens of the peptide's terminal carboxylic amide group and of Ala229 of the opposing protomer (Figure 1.5-8B). This induced an altered conformation of the loop linking helix H3 and the last helix, H4, which was in turn not observed in the structure. These differences may be causative of the entirely different crystal packing of the complex (Figure 1.5-8C,D).

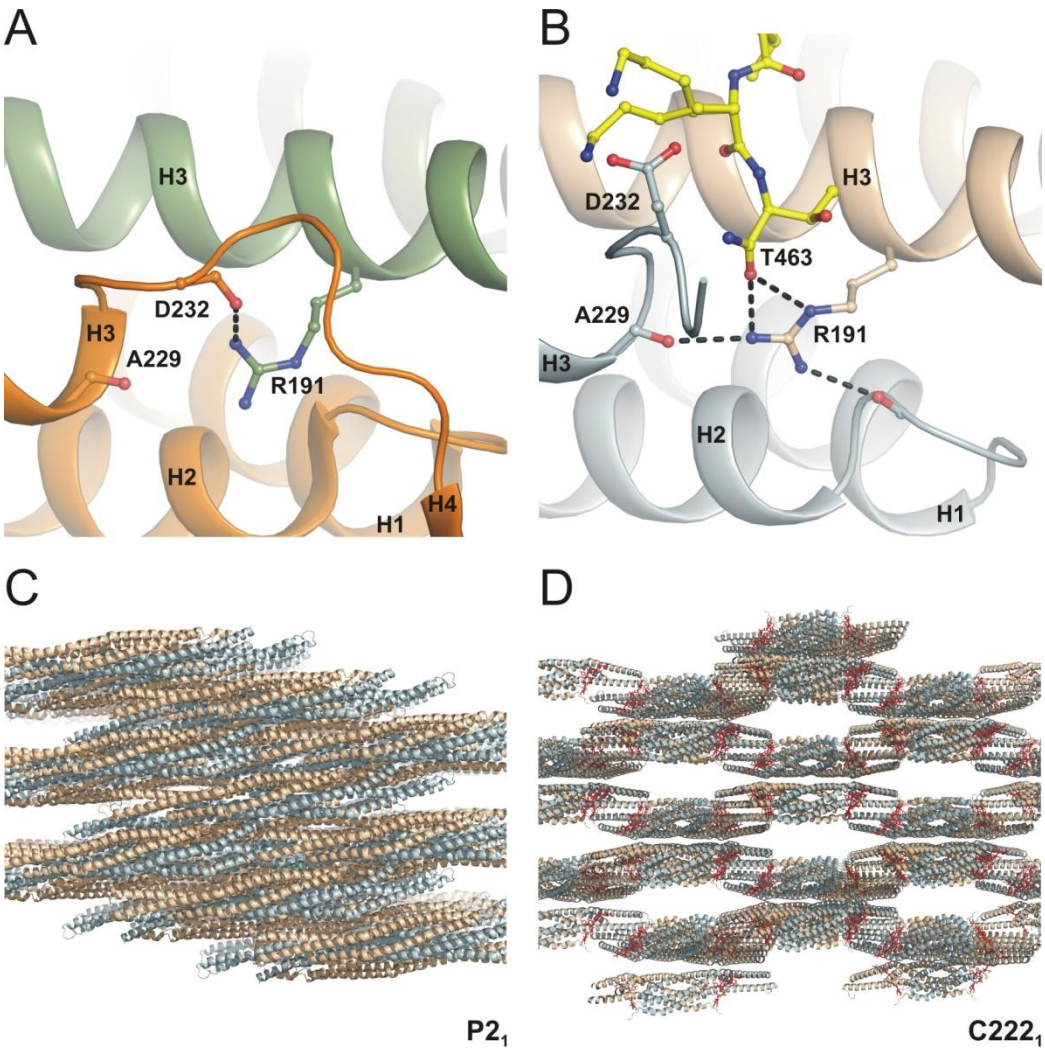

Figure 1.5-8: Differences in crystal packing of free and peptide loaded IMD$_{IRSp53}$ facilitates a conformational change of H4. (A) In the free structure (PDB 1Y2O), Arg191 of one monomer (green) binds to the loop linking helix H3 and helix H4 of the other monomer (orange) and thereby stabilises helix H4 adjacent to the dimer. In the structure of the complex (B), the peptide's terminal amide group (yellow) binds to the guanidinium group of Arg191 of one IMD monomer (wheat), leading to a conformational change of helix H4 and the preceding loop of the other monomer (blue). (C) Crystal packing according to Millard *et al.* of the free IMD (PDB 1Y2O). IMDs are densely packed. The two protomers of each dimer are in cyan and wheat, respectively. (D) Crystal packing of the peptide loaded IMD displays a more loose array, explaining the undetectability of the loop between H2 and H3 and of entire H4. Tir peptides are shown in red. Space groups are indicated.

1.5.4.3 The Tir binding site is conserved

The paralogues IRSp53 and IRTKS interact with both EHEC effectors Tir and $EspF_U$ *via* their IMD and SH3 domains, respectively. In contrast, the third family member in mammals, FLJ22582, only binds $EspF_U$ (Weiss *et al.*, 2009). The sequence identity for the IMDs of IRSp53 and IRTKS is 47% and all residues contacting the NPY motif are conserved (Figure 1.5-9).

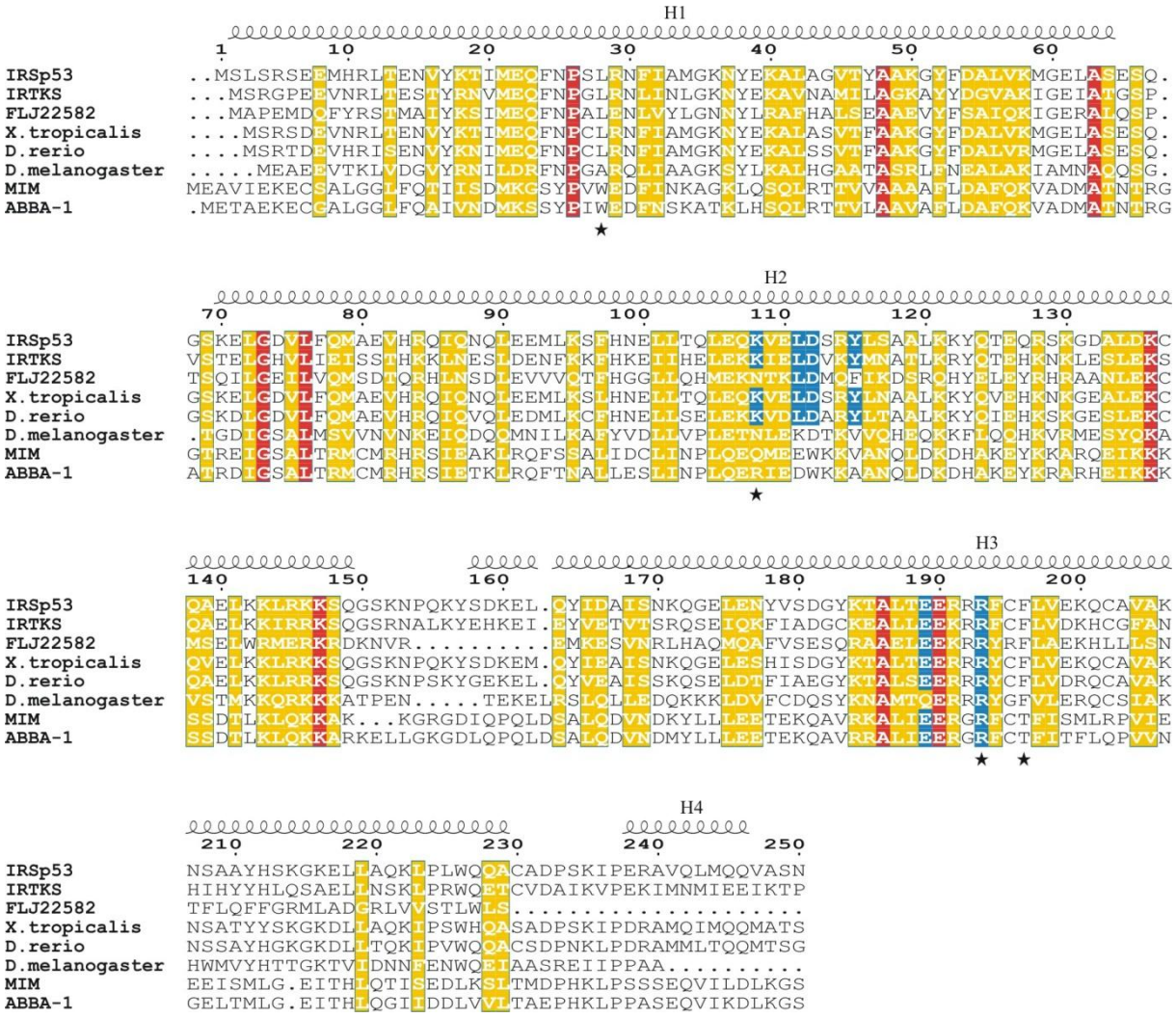

Figure 1.5-9: IMD sequence conservation. Sequence alignment of the known human IMD domains of IRSp53, IRTKS, FLJ22582, MIM and ABBA-1 and IRSp53 orthologs from *Xenopus tropicalis* (claw frog), *Danio rerio* (zebrafish) and *Drosophila melanogaster* (fruit fly) (GenBank NP_059344, NP_061330, NP_079321, NP_055566, NP_612392, NP_001121436, NP_001035335, NP_729679). Blue boxes indicate IRSp53 residues in contact to the Tir NPY motif. Red and yellow boxes highlight identical and similar residues, respectively. Positions that were mutated in this study are marked by asterisks.

The IMD of FLJ22582, however, has only a sequence identity of 28% to IRSp53, and the residues forming the binding pocket differ markedly. The most prominent difference is FLJ22582 Asn108, which replaces Lys108 of IRSp53 (Figure 1.5-10). Lys108 is one of most important residues in the binding pocket, stabilising Tyr458 of the NPY motif (Figure 1.5-6B). The lack of Tir binding in FLJ22582 may thus be explained by its absence. The Tir binding site is highly conserved within the group of IRSp53 and IRTKS orthologues of evolutionary distant vertebrates including claw frog and zebrafish, but not in invertebrate orthologues. The site is also not conserved in the human IMD proteins MIM and ABBA-1 (Figure 1.5-9).

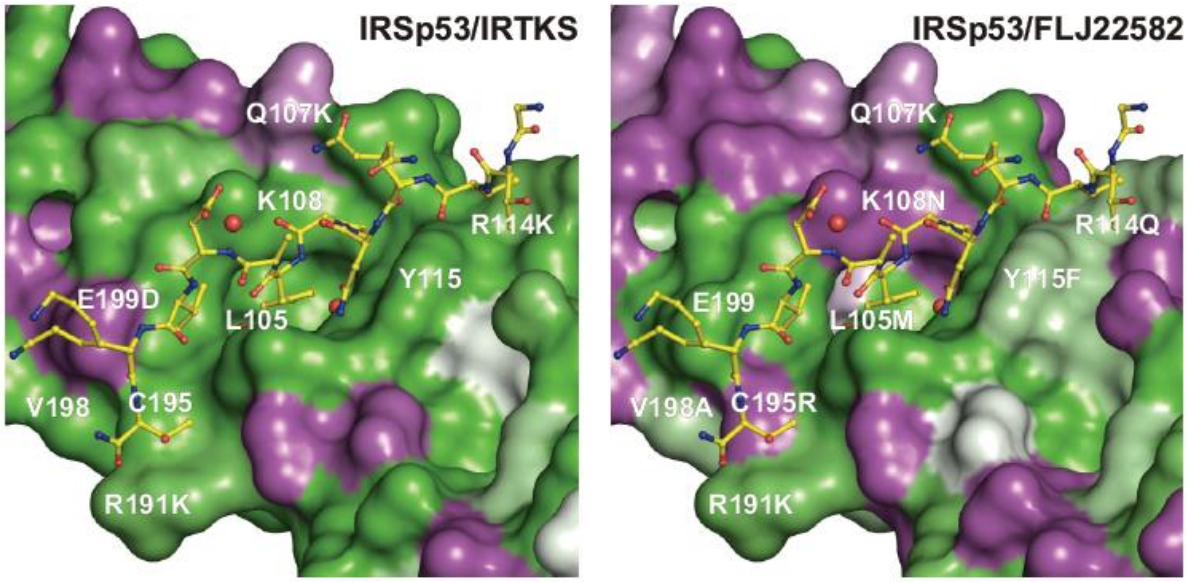

Figure 1.5-10: Mapped IMD sequence conservation. Sequence conservation between the IRSp53 residues, constituting the hydrophobic binding pocket, and IRTKS (left) or FLJ22582 (right). Deep green indicates identity, while purple represents differing residues.

1.5.4.4 ITC analysis to study the effect of mutations on the IMD-IRSp53:Tir complex formation

The complex formation between the IMD_{IRSp53} and the Tir-derived peptide was studied by isothermal titration calorimetry (ITC). The stoichiometry of two peptide molecules per IMD homodimer was confirmed, and binding was apparently non-cooperative. A dissociation constant of 59 ± 10 µM was determined (Figure 1.5-11), which indicates a weak interaction, comparable to SH3-domain interactions with proline-rich motifs (Ball *et al.*, 2005; Kuriyan and Cowburn, 1997; Mayer and Saksela, 2004). A corresponding 12-mer peptide was derived from EPEC-Tir (SEVV**NPY**AEVGG) and a similar dissociation constant of 37 ± 14 µM was obtained.

Binding of the IMD_{IRSp53} to additional peptides with motifs related to NPY was studied by ITC, including the ^{456}NPY^{458}→NPF mutant, described in section 1.3.6, and a peptide derived from the insulin receptor comprising the related NPXY motif. No binding was detectable by ITC for these NPF and NPXY motif containing peptides (data not shown), demonstrating preference of IMD for the NPY motif. To interpret these results, it should be noted that the protein concentration used for ITC (100 µM) was close to the dissociation constant of the wild type domain and the Tir-derived peptide (60 µM), and it would not have been possible to detect binding at lower affinities.

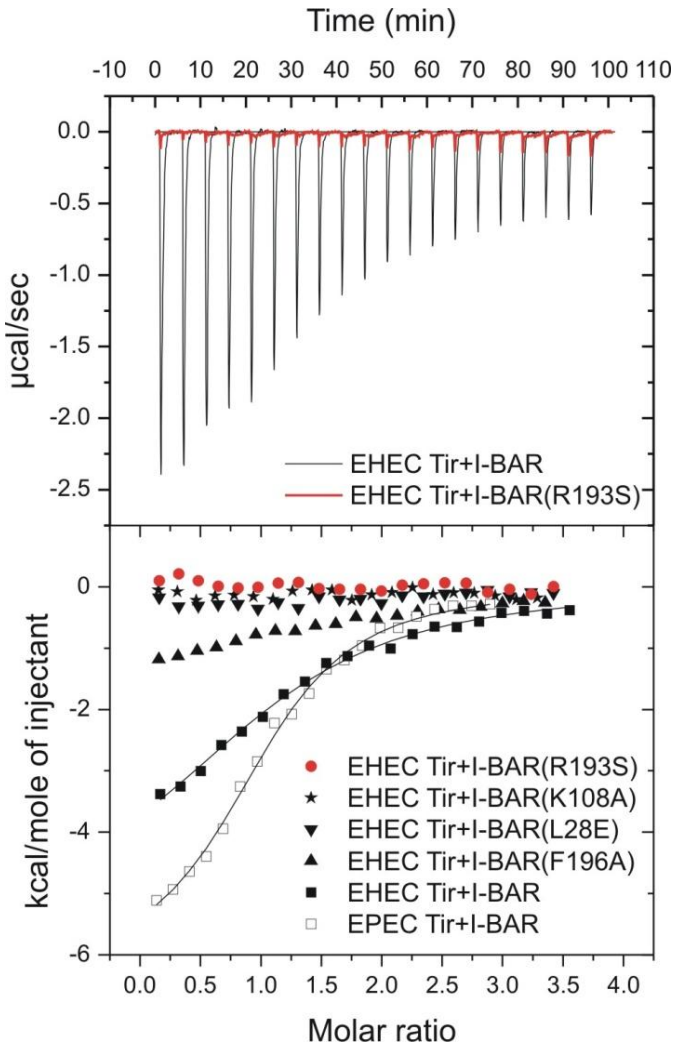

Figure 1.5-11: Isothermal titration calorimetry of the binding reaction of IMD_{IRSp53} and four IMD point mutants with peptides derived from EHEC and EPEC Tir, comprising the NPY motif. The upper panel shows considerable heat release upon titration of the wild type IMD_{IRSp53} (black) with the EHEC Tir-derived peptide, in contrast to the R193S mutant (red). The heat data, corrected for dilution, is plotted against the molar ratio of peptide molecules and IMD protomers in the lower panel. Sigmoidal curves typical of exothermic binding reactions were obtained for the wild type IMD_{IRSp53} and the mutant F196A, but not for the R193S, K108A and L28E mutants.

Close inspection of the binding site allows the prediction of IMD point mutants, impaired in Tir binding without disturbing the overall IMD structure (Figure 1.5-9). These point mutants were generated in-house by the group of Theresia Stradal. Substitution of Leu28 by glutamic acid (L28E) was predicted to fill the hydrophobic pocket. The residue Lys108 forms a hydrogen bond and hydrophobic contacts with the peptide. It is changed to an asparagine in FLJ22582, the IRSp53 paralogue that does not recognise Tir. The conformation of the Lys108 side-chain is stabilised by Phe196. Arg193 may be critically important for recognizing Tir, since it is the residue with most interactions with the peptide (Figure 1.5-7) and since the salt bridge to Asp112 appears to stabilise the binding pocket. No binding of the Tir-derived peptide was detected by ITC when Leu28, Lys108, or Arg193 were substituted in IMD_{IRSp53} (L28E, K108A, R193S in Figure 1.5-11). Binding of the Phe196 point mutant F196A to the peptide was weaker than for the wild type domain but still clearly detectable.

1.6 Discussion

1.6.1 The NPY motif of Tir: a 'hot spot' for IMD recognition

The crystal structure presented here reveals the molecular details of the interaction between EHEC Tir and the IMD of human IRSp53. It elucidates why the highly conserved NPY motif of EHEC Tir is critical for interaction with the IMD of IRSp53 and IRTKS, whereas neighbouring residues can be substituted with alanine without any effect (Brady *et al.*, 2007; Weiss *et al.*, 2009). The specificity of recognition of Tir by the IMD is determined by the characteristic features of the NPY motif's three residues, 456-458. The Asn456 side chain is bound by specific hydrogen bonds. The Tyr458 phenyl ring projects into a hydrophobic pocket, forming extensive hydrophobic interactions to the IMD, while the phenolic hydroxyl group is bound by two water molecules at the bottom of the pocket. Correspondingly, substitution of Tyr458 with alanine virtually abolished pedestal formation by EHEC bacteria, while substitution with phenylalanine led to a milder defect (Brady *et al.*, 2007). The pyrrolidine ring of Pro457 of the NPY motif fits into a shallow hydrophobic cavity on the IMD surface. It rigidifies the backbone and supports the formation of two overlapping type I β-turns in the Tir polypeptide chain.

The structure described here adds the IMD to the list of known peptide binding domains. The binding energy of protein-protein and peptide-protein interactions generally depends on only a small number of 'hot spot' residues (London *et al.*, 2010), obviously represented here by the NPY motif, including the tightly bound Tyr458. Interestingly, tyrosine ranks on third position of the most frequent amino acid residues identified in peptide-protein complexes (London *et al.*, 2010). Specific recognition of phosphorylated Tyr474 by Nck proteins is essential for pedestal formation by EPEC, as mentioned above. Other tyrosines of EPEC Tir at positions 483 and 511 also become phosphorylated and are essential for recruiting SHIP2, a host inositol phosphatase controlling pedestal formation by generating a phosphatidylinositol 3,4-bisphosphate-enriched lipid platform (Smith *et al.*, 2010). It appears that tyrosine-mediated interactions of Tir with specific peptide recognition domains of the host cell are critical for pedestal formation of both EHEC and EPEC.

1.6.2 The conserved Tir binding pocket of IMD: the essential residues and their relevance for NPY recognition

The crucial role of certain residues within the IMD NPY binding site revealed by the structure was confirmed by site-directed mutagenesis and isothermal titration calorimetry (section 1.5.4.4). Further binding studies as pull-down assays from lysates and interaction assays in living cells, using the Protein Interaction Platform (PIP) assay in yeast (Schmitz *et al.*, 2009),

were carried out in cooperation with the group of Prof. Dr. Theresia Stradal (Institute for Molecular Cell Biology, University of Münster, Germany), Dr. John M. Leong (Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, MA, USA) and Cammie F. Lesser (Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA, USA), respectively.

Table 1.6-1: IMD mutants used to study the interaction with Tir. ITC and pull-down assays were carried out using the IMD of IRSp53, whereas the IMD of IRTKS was used in the PIP assay. The sequence of IMD_{IRTKS} aligns the sequence of IMD_{IRSp53} with a shift of -1. X indicates the use of the domain in the corresponding assay.

IMD_{IRSp53}	IMD_{IRTKS}	ITC	Pull-down assay	PIP assay
Wt	Wt	X	X	X
L28E	L27E	X	X	X
K108A	K107A	X	X	X
R193S	R192S	X	X	X
F196A	F195A	X	X	X
K108A/R193S	K107A/R192S		X	X
	K107A/F195A			X
R193S/F196A	R192S/F195A		X	X

The pull-down assays were used to probe the ability of the described IMD mutants (section 1.5.4.4 and Table 1.6-1) to interact with the Tir-derived peptides or, in contrast to the isothermal titration calorimetry, with the entire cytoplasmic carboxy-terminal tail of EPEC- and EHEC-Tir (TirC). The pull-down assays were performed from lysates of COS7 cells, ectopically expressing the IMD variants as amino-terminal GFP fusions, using biotinylated synthetic Tir-derived peptides immobilised on streptavidin-agarose beads as bait. Under these conditions, wild type EPEC and EHEC-derived peptides showed significant binding to the IMD of IRSp53, and could effectively deplete GFP-tagged IMDs from the lysate and enrich in the precipitate (Figure 1.6-1). In contrast, neither the IMD mutants K108A, F196A (Figure 1.6-1), nor L28E, R193S, or the IMD double mutants K108A/R193S and K108A/F196A (not shown) could be detected in the precipitates. *Vice versa* experiments were performed by using recombinant His-tagged wild type and mutant IMDs as baits in pull-down assays from lysates of COS7 cells, ectopically expressing the entire TirC. The recombinant wild type IMD effectively bound to TirC from both EPEC and EHEC. In accordance with the data obtained by ITC, weak but reproducible binding of TirC was also

observed for the F196A mutant of IMD_{IRSp53}, while all other IMD mutants were abrogated for binding (Figure 1.5-11 and Figure 1.6-1B).

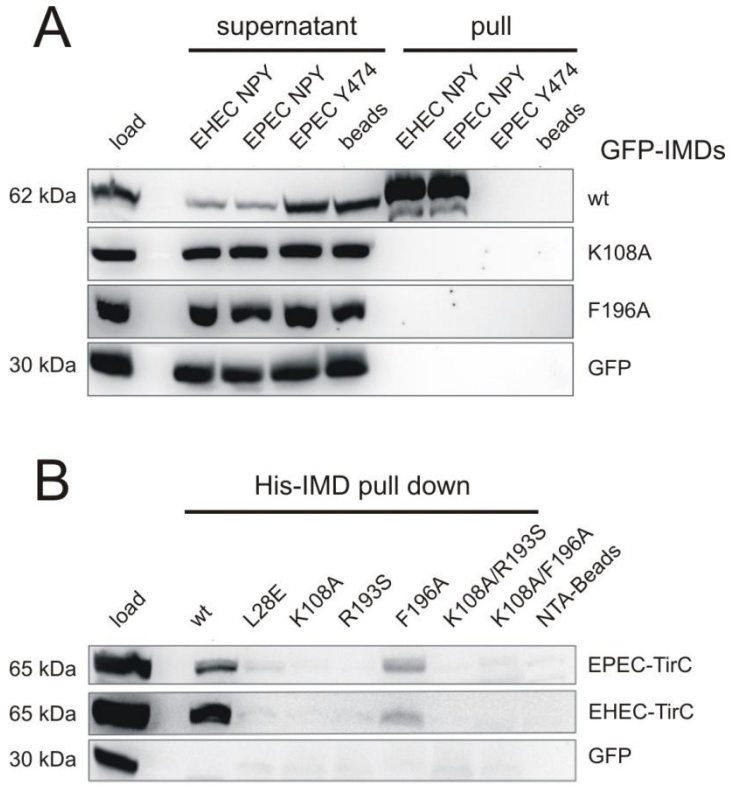

Figure 1.6-1: Pull-down assays of Tir with wild type and mutant IMD of IRSp53. (A) Pull-down assays of Tir-derived peptides from lysates of mammalian cells ectopically expressing GFP-tagged IMD variants of human IRSp53 as indicated on the right. Shown are cell lysates before (load) and after (supernatant) pull-downs and precipitates (pull) of three different peptides as indicated above. EHEC Tir-derived peptides that bind to Nck instead of IRSp53 (Y474) as well as streptavidin beads were employed as bait-control and GFP alone in the cell lysate was the prey-control. (B) Pull-down assays of recombinant His-tagged IMD from lysates of COS7 cells ectopically expressing the cytoplasmic carboxy-termini of EPEC or EHEC Tir as indicated on the right. Shown is the cell lysate before the pull-down (load) and the precipitates (His-IMD pull-down) of wild type IMD (wt) and six variants as indicated above. Note the weak binding to mutant F196A in accordance with the ITC measurements.

PIP-assays were used to investigate whether the mutations analysed above within the IMD of IRSp53 would also affect the interaction of Tir with IRTKS. Given the strong conservation of the Tir binding site, IRSp53 and IRTKS most likely bind to the Tir receptor in the same fashion. EHEC TirC and a ^{456}NPY^{458}→AAA substitution mutant (TirC AAA) (Vingadassalom *et al.*, 2009) were fused to the scaffolding protein µNS that forms large focal inclusions ('platforms') in yeast cells. Co-expression of GFP-tagged IMD of IRTKS with the TirC fusion protein resulted in complex assembly and formation of fluorescent foci in almost all cells (Figure 1.6-2A). The TirC AAA mutant was not able to alter the diffuse intracellular localisation of IMD_{IRTKS}, as expected. Substitution of IRTKS Lys107 or Phe195

(corresponding to IRSp53 Lys108 and Phe196) partially inhibited the interaction with TirC (K107A and F195A, Figure 1.6-2A,B). This was exacerbated by combining both mutations (K107A/F195A). Strong inhibition was also observed for the substitution of Leu27 by glutamic acid (L27E). Finally, substitution of Arg192 in IRTKS completely abolished the Tir:IMD_{IRTKS} interaction (R192S), as the double mutations R192S/F195A and K107A/R192S did (Figure 1.6-2A,B).

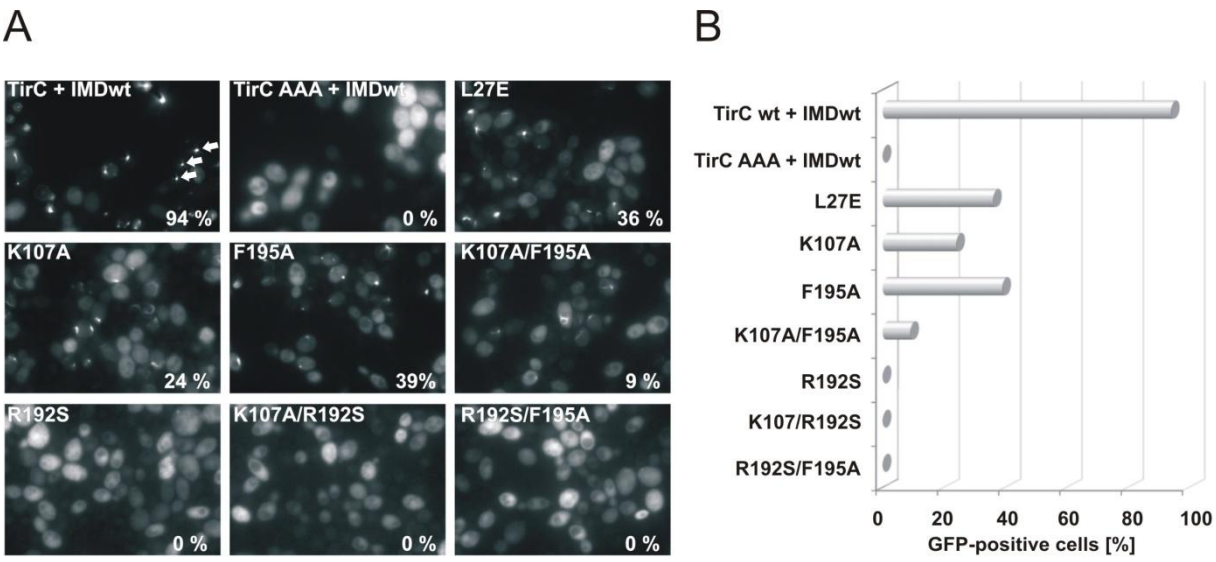

Figure 1.6-2: Protein Interaction Platform assay of Tir with wild type and mutant IMD of IRTKS *in vivo*. (A) Yeast cells co-expressing µNS-TirC and wild type or mutant GFP-tagged IMD of IRTKS were visualized by fluorescence microscopy 4-5 hours post-induction of fusion protein expression. Exemplary GFP inclusions are marked by arrows. The percentage of cells containing GFP inclusions was determined visually and is indicated at the bottom right of each image and (B) visualized in the histogram.

The interaction of the IMD of IRSp53 and IRTKS with Tir *in vitro* and in living cells was thus impaired by substitution of residues constituting the peptide binding site (Figure 1.5-11, Figure 1.6-1 and Figure 1.6-2). The results were conclusive, also concerning residual binding that was always found to be strongest for the F196A or F195A mutants, respectively. The tight interaction of Arg193 with Tir in the crystal structure was reflected by the complete loss of binding upon substitution of the corresponding residues in IRSp53 and IRTKS. Substitution of Lys107 or Lys108 in respectively IRTKS or IRSp53 considerably reduced or abolished binding. The homologous FLJ22582 has an asparagine at this position, which is probably the reason for its lack of binding to Tir (Weiss *et al.*, 2009).

Full-length IRSp53 was shown to equally well precipitate with both EPEC and EHEC Tir-derived peptides, containing this NPY motif (Weiss *et al.*, 2009). Moreover, both GFP-tagged IRSp53 and endogenous IRSp53 can be readily detected in both EPEC and EHEC pedestals

(Weiss *et al.*, 2009). However, in contrast to EHEC, that essentially require IRSp53, it is dispensable for EPEC pedestal formation as exemplified by using IRSp53 knockout cells (Weiss *et al.*, 2009). Notwithstanding this, introduction of $EspF_U$ into EPEC drastically increases Nck-independant pedestal formation (Brady *et al.*, 2007). Thus, IRSp53 effectively binds to EPEC Tir but no information is available to date on the state of EPEC Tir Y454 phosphorylation, when bound to IRSp53.

In Summary, our current data confirm a virtually complete conservation of the binding characteristics between IRSp53 and IRTKS to the carboxy-termini of EPEC and EHEC Tir.

1.6.3 The conserved Tir binding pocket of IMD and the preference for NPY motifs

The mode of binding of the NPY motif is strikingly similar to the recognition of NPXY and NPF motifs by phosphotyrosine binding (PTB) and Eps15 homology (EH) domains, respectively. The NPXY and NPF motifs also form stabilised Asn-Pro type I β-turns, their asparagine side-chains are specifically hydrogen bonded and their aromatic side chains enter hydrophobic pockets in the respective domains (de Beer *et al.*, 2000; Eck *et al.*, 1996). Binding of peptides containing an NPXY or NPF motif by IMD was not detected by ITC, indicating preference for NPY over related motifs.

1.6.4 Tir is the first unambiguously IMD binder

The striking specificity of the recognition of the NPY motif by the IMD suggests that the domain has evolved to recognise this motif, a feature that just becomes usurped by pathogenic *E. coli*. Thus it is tempting to speculate that this binding site may have cellular ligands that harbour NPY motifs, which occur in up to 3% of human proteins. However, up until now, no binding partners for the IMD domains of IRSp53 and IRTKS other than Tir have been pinned down unambiguously because the *in vivo* relevance of its binding to both Rac1 (Disanza *et al.*, 2006; Krugmann *et al.*, 2001; Miki *et al.*, 2000) and actin (Mattila *et al.*, 2007; Millard *et al.*, 2005) is still the subject of debate.

1.6.5 The affinity of Tir

The almost parallel orientation and the distance of the two Tir-derived peptides in the complex suggest that one IRSp53 homodimer can link two Tir molecules *in vivo*. An affinity of

60 µM was found for the binding reaction under *in vitro* conditions, which is comparable to the relatively weak interactions between SH3-domains and proline-rich motifs (Ball *et al.*, 2005; Kuriyan and Cowburn, 1997). However, the affinity and specificity of ligand binding by SH3 can be greatly enhanced by additional contacts between the SH3 loop regions and residues of the ligand outside the proline-rich motif (Arold *et al.*, 1998; Mayer and Saksela, 2004). This is greatly illustrated by the SH3 blocker of Kardinal and colleagues, optimizing a high affinity SH3 blocker peptide with a K_D in the nanomolar range that competes for the SH3 of CRKL with Bcr-Abl and hence reduces the proliferation of primary blast cells from chronic myeloid leukemia patients (Kardinal *et al.*, 2000). Thus, regions outside the Tir-derived peptide might likewise contribute to the binding to IRSp53. This notion is supported by the fact that weak binding of the IMD mutant F196A to a longer carboxy-terminal Tir fragment was observed in pull down experiments (Figure 1.6-1B), while binding to short Tir-derived peptides was not detectable (Figure 1.6-1A). Furthermore, Tir forms homodimers and is clustered beneath the attached bacteria upon interaction with intimin (Luo *et al.*, 2000; Thierry *et al.*, 2004). Together with the presence of two binding sites on the IMD homodimer, this is expected to lead to a higher apparent affinity due to the avidity effect and, consequently, efficient recruitment of IRSp53 at physiological concentrations.

1.6.6 Model of the intimin:Tir:IRSp53 complex

The EHEC bacterium utilises a large protein scaffold including bacterial and host proteins that immobilises the bacterium and leads to N-WASP mediated Arp2/3-complex dependant actin polymerisation (Weiss *et al.*, 2009). Figure 1.6-3 shows a model of this scaffold consisting of Tir, intimin and IRSp53. An alternative model with one IRSp53 homodimer linking two distinct Tir homodimers appears also possible. IMDs recognise and even induce convex deformation of lipid membranes (Scita *et al.*, 2008; Suetsugu *et al.*, 2006). Membrane binding by human IRSp53 is short-lived and highly dynamic (Saarikangas *et al.*, 2009) and it is dispensable for pedestal formation (Weiss *et al.*, 2009). The orientation of the Tir-derived peptides on the IMD suggest that the flat IMD surface, which is assumed to contact the plasma membrane, is oriented towards the membrane in the complex, thus supporting the possibility of Tir-bound IRSp53 contacting the plasma membrane.

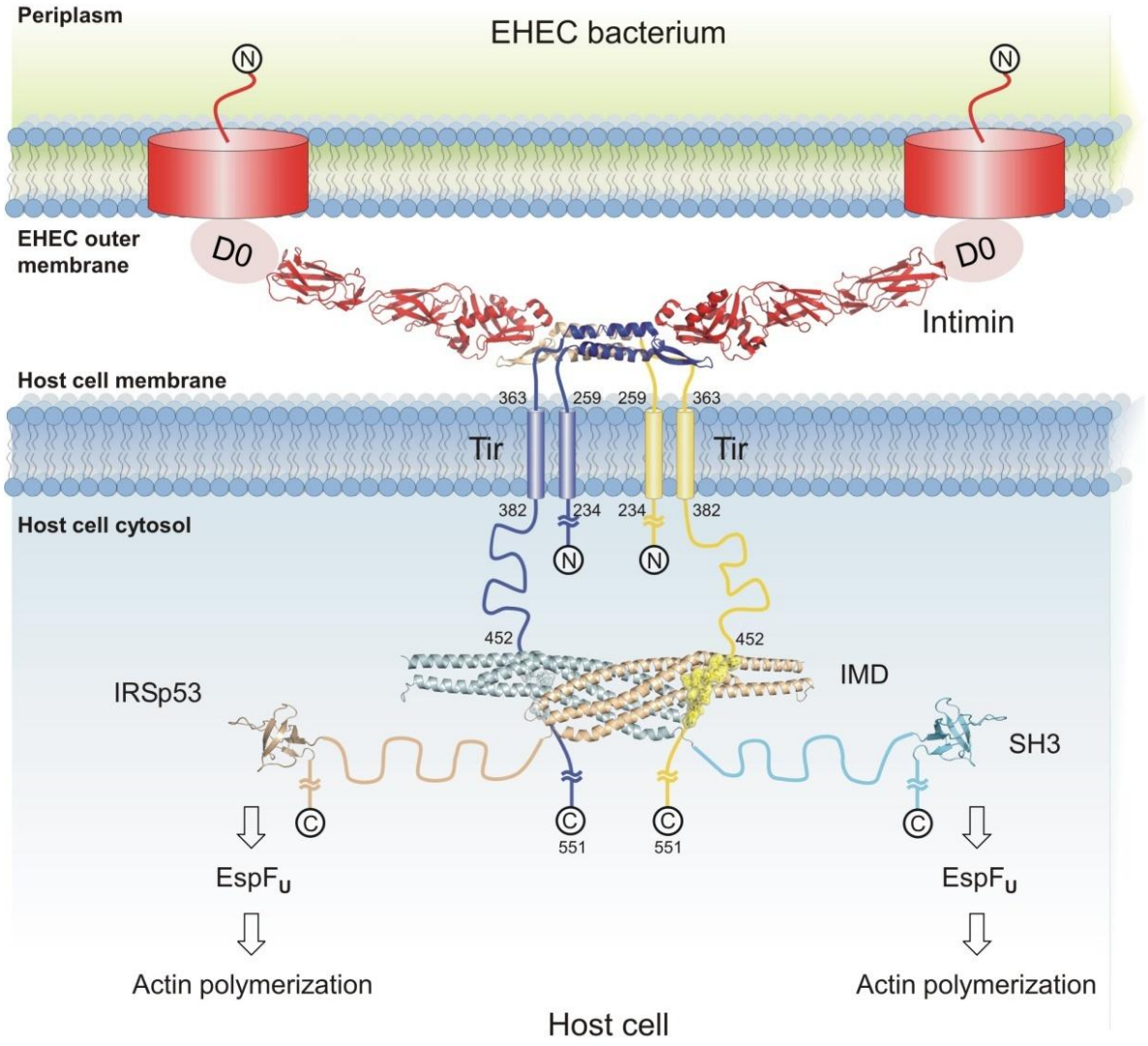

Figure 1.6-3: Model of the intimin:Tir:IRSp53 complex. A model of the scaffold formed by EHEC Tir, intimin, and human IRSp53 was drawn to scale based on the available structures. Whether the Tir-bound IMD interacts with the plasma membrane directly is currently unknown. The extracellular Tir and intimin domains are represented by the structure of the complex of the homologues EPEC proteins by Luo *et al.* (PDB entry 1FO2). The model excludes the amino-terminus of Tir, which modulates pedestal length (Campellone *et al.*, 2006). The IRSp53 SH3 domain is represented by the similar PDB entry 2KXC. Regions of unknown structure are represented by ropes with end-to-end distances corresponding to root mean square end-to-end distances of random coils of the respective chain length (Tanford *et al.*, 1966).

1.7 Outlook

1.7.1 NPY motif containing host proteins may represent physiological ligands for IRSp53/IRTKS

The emergency of the critically increasing incidence of antibiotic resistance due to the widespread and, to some extent, careless use of antibiotics requires a more detailed understanding of the molecular basis of host pathogen interactions to generate alternative methods to treat infectious diseases. Moreover, diseases like the HUS, described to be triggered by the use of antimicrobial substances (see section 1.1.4.2), currently lack a well-established method of treatment and rely on the development of such alternative strategies. Structural insights into the key mechanisms of infection of human pathogenic *E. coli* can pave the way to design inhibitors to disrupt key interactions during infection and enable the treatment of infectious diseases without the use antibiotics.

Several studies already aim at the development of strategies to neutralise Shiga toxins or to disrupt the mechanism of adhesion of human pathogenic bacteria like blocking the interaction of Tir:intimin of intimin positive EHEC. However, the Tir:IRSp53 interaction currently seems to be no drug target to treat EHEC infections as this interaction seems to trigger only actin polymerisation. Inhibition of actin polymerisation and thus pedestal formation after EHEC infection might, to some extent, probably improve absorption and reduce intestinal tissue damage and the dissemination of EHEC, which is suggested to be enhanced by the pedestals which move along the cell surface, the so-called ‘pathogen surfing’. But, according to the current knowledge, inhibited pedestal formation will probably not affect adhesion or toxin production of EHEC, the central issue of EHEC infections. Nevertheless, unveiling the molecular mechanism of pedestal formation and the investigation of the bacterial benefit of this feature will contribute to the substantial understanding of EHEC infections and hence the development of strategies to treat this disease.

The data of this study characterise the NPY motif of Tir as a ‘hot spot’ for the interaction with the host’s IMD. The domain seems to be evolutionary evolved to recognise this motif, a feature that becomes usurped by pathogenic *E. coli*, and it is tempting to speculate that this binding site may have cellular ligands that harbour NPY motifs, which occur in up to 3% of human proteins. The huge variety of possible ligands to this binding site of the host’s IMD illustrate again, that even if the IMD:Tir interaction represented a pharmacological relevant drug target, which is currently not the case, the Tir binding pocket of IMD seems to be inappropriate due to the presumably adverse effects by blocking also diverse cellular ligands. But interestingly, among these NPY motif containing proteins, numerous candidates can be found that are related to cellular processes that may involve IRSp53 or IRTKS (see Table 1.7-1). Thus the discovery and characterisation of this new binding pocket may lead to the

identification of new interaction partners, possibly involved in the pathogenesis and cytoskeletal regulation. Furthermore, these studies will also contribute to a better understanding of the cellular role of IRSp53. It will be an exciting task to probe these interactions and analyse the recently described IRSp53 knockout mouse (Sawallisch *et al.*, 2009) for potential phenotypes that may relate to defective IMD:NPY interactions.

Table 1.7-1: Potential NPY-containing interactors of IMD proteins. The SwissProt database comprises 20,329 human proteins, of which 749 contain the NPY motif as detected using Scansite (http://scansite.mit.edu/dbsequence_one.html). Candidates that potentially might interact with the IMD of IRSp53 or IRTKS were selected according to the following criteria: (i) nuclear, intraluminal or extracellular motifs are excluded; (ii) the motif is conserved at least between human and mouse orthologs and (iii) conservation between paralogues is preferred. The selection shown is further confined based on involvement in IRSp53-signalling (e.g. Shank), in cytoskeleton reorganisation (e.g. LIMK), in Rho-GTPase signalling (Intersectin), and for those that were implicated in filopodium (or similar like stereocilia/microvilli, e.g. Myosin X) formation before (de Groot *et al.*, 2011).

Short name	Full name	NPY position	NPY location	Protein class
PK3CB	Pi3K, p110b, Phosphtidylinositol-3 kinase, catalytic subunit	482-84	C2 domain	phospholipid kinase
ROCK2	Rho-associated protein kinase 2	1195-97	very C-terminus	protein kinases
LIMK1 & -2	LIM domain kinase 1	512-14	kinase domain	
ITSN1 & -2	Intersectin 1 (Cdc42 GEF)	1618-20	C2 domain	regulators of Rho-GTPase activity
ARHGAP10	GRAF2 GAP10 associated with FAK (RhoA, Cdc42)	625-27	between Rho-GAP and SH3	
Myo IA	Brush border myosin 1 /BBM-I	49-51	head domain	myosins
Myo X	Myosin 10 (filopodia)	104-06	head domain	
MyoXV	Myosin 15 (stereocilia)	423-25 1263-65	head domain NT head domain CT	
CLAP1 & -2	CLIP115 associated protein1	1265-67	CLIP-binding site	microtubule tip proteins
SHANK2	SH3 and multiple ankyrin repeat domains protein 2	608-10	central part	partners of IRSp53
BAI2	Brain-specific inhibitor of angiogenesis 2 (GPCR)	1417-19	cytoplasmic tail	

1.7.2 The Tir:IMD interaction is likely to compete with membrane binding

The cigar shaped IMD dimer possesses positively charged amino acids exclusively concentrated at the ends of IMD, interacting with high affinity to $PI(4,5)P_2$, and localise the IMD to the plasma membrane (Mattila *et al.*, 2007; Saarikangas *et al.*, 2009). However, it is also shown that membrane binding is not required for the formation of pedestals by EHEC bacteria (Weiss *et al.*, 2009). It would be interesting to examine whether the carboxy-terminal region of Tir competes with the membrane for IRSp53 binding or if membrane bound IRSp53 is also capable to interact simultaneously with this Tir fragment. This could be examined by e.g. vesicle cosedimentation assays (Saarikangas *et al.*, 2009), where the ability of labeled Tir to cosediment with membrane bound IMD would be examined.

1.7.3 The binding affinity of Tir to IMD

The quite low dissociation constant of 60 µM of the EHEC Tir:IMD interaction, measured by ITC, may be due to the use of a short peptide instead of the whole caboxy-terminal cytoplasmic Tir. As already discussed (section 1.6.5), the SH3 domain of IRSp53 as well as the avidity effect may contribute to this interaction and possibly dramatically increase the binding affinity. Thus, it would be exciting to test whether other regions of Tir and IRSp53 also contribute to this interaction. Furthermore, it would be of interest to test the effects of IRSp53 mutants in physiological relevant cell culture models, as the pull down assays were performed using cell lysates and the PIP assays were performed in yeast.

2 Chapter II

Structural Characterisation of Amorfrutins Bound to the Peroxisome Proliferator-activated Receptor gamma

2.1 Introduction

2.1.1 The nuclear receptor PPARγ

The peroxisome proliferator-activated receptor gamma (PPARγ) belongs to the so-called adopted orphan receptors of the nuclear receptor superfamily. The name is derived from the ability of the first identified member of this family, PPARα, responding, in contrast to the other family members, to various compounds that induce peroxisome proliferation (Tontonoz and Spiegelman, 2008). The PPARs exhibit distinct tissue distributions with different activities. PPARα is expressed in liver, kidney, heart, and muscle where it regulates fatty acid catabolism. PPARδ is ubiquitously expressed with less defined functions, but is implicated in e.g. wound healing and keratinocyte differentiation (Michalik *et al.*, 2001). PPARγ, the target protein of this study, is mainly expressed in adipocytes and macrophages. PPARγ consists of an amino-terminal, ligand-independant activation function 1 (AF1), a DNA binding domain which is highly conserved among the superfamily, followed by a hinge region and a carboxy-terminal conserved ligand-binding domain (LBD), containing the activation function 2 (AF2) (Figure 2.1-1A) (Evans, 1988; Glass and Ogawa, 2006). PPARγ acts in the nucleus as a heterodimer with the retinoid X receptor RXRα and is a ligand-activated transcription factor (Berger and Moller, 2002), interacting with food ingredients like fatty acids and their metabolites (Figure 2.1-1B) (Kliewer *et al.*, 1997; Tontonoz and Spiegelman, 2008). PPARγ is a well-known drug target of e.g. type II diabetes treatment as it works as a dominant sensor and regulator in adipose cell differentiation and glucose and lipid metabolism, but

plays also a key role in inflammation, atherosclerosis and cancer, and has strong anti-growth properties (Murphy and Holder, 2000; Tontonoz and Spiegelman, 2008).

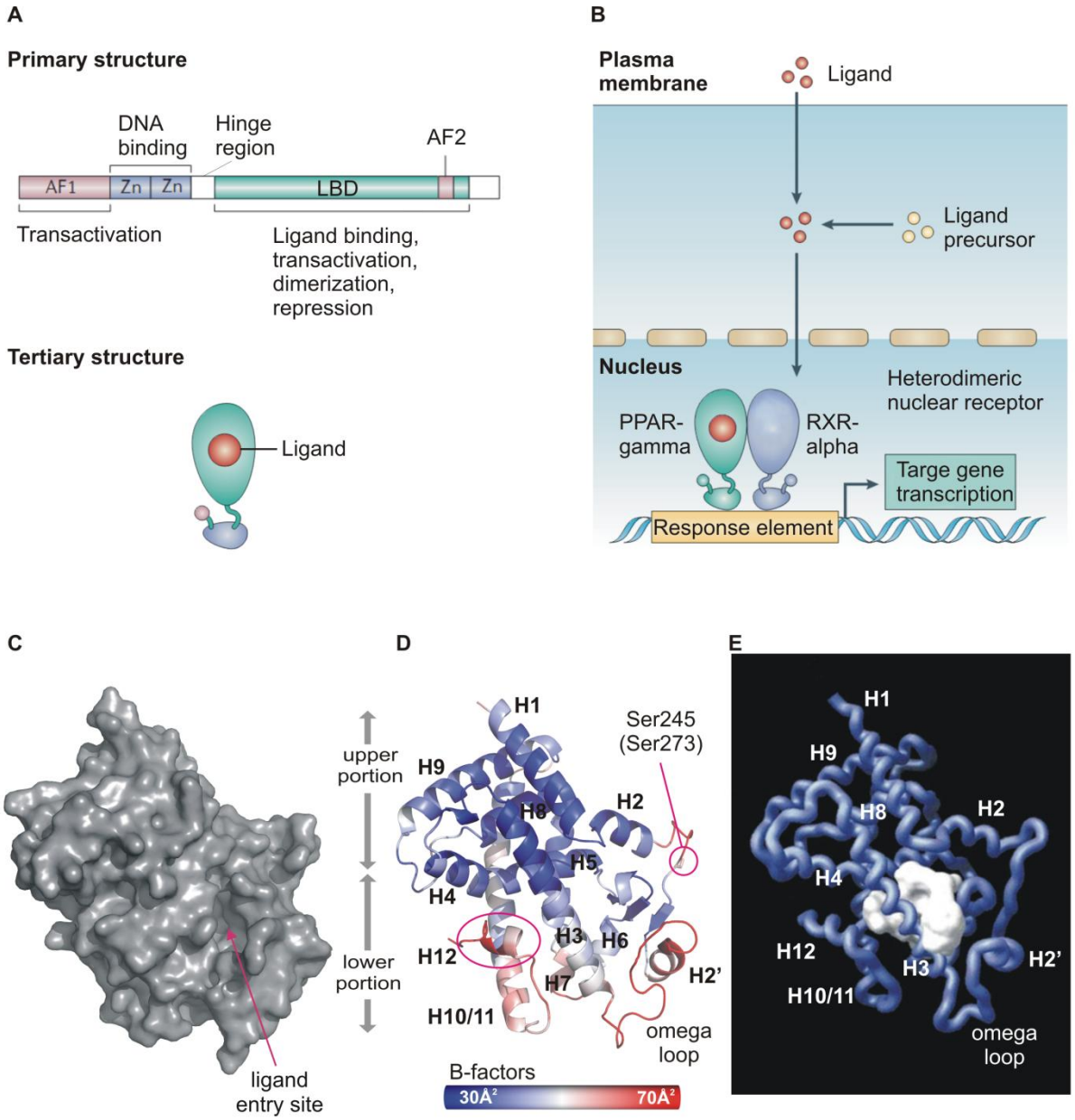

Figure 2.1-1: Domain organisation of PPARγ and the structural characteristics of its LBD. (A) PPARγ consists of an amino-terminal activation function domain 1 (AF1), a central DNA-binding domain, and a carboxy-terminal ligand-binding domain (LBD) (Glass and Ogawa, 2006). (B) PPARγ and the mechanism of DNA binding according to Glass and Ogawa (2006). PPARγ heterodimerises with RXRα and binds to DNA in a ligand-dependant manner. (C-E) The protein surface and ribbon representation of PPARγ's LBD reveal a huge ligand entry site in the lower portion of the protein, located between helix H3 and H2'. (D) The receptor dynamic is illustrated by plotting the temperature factors (B-factors) on the structure (PDB 1PRG), revealing the omega loop and helix H12 as the most mobile segments of the LBD. The phosphorylation site at Ser245 (Ser273 in PPARγ2), and helix H12 are highlighted. (E) The huge ligand binding cavity of PPARγ is illustrated by white surfaces (Nolte *et al.*, 1998).

2.1.1.1 The activation function 1 (AF1)

The amino-terminal AF1 of the nuclear receptor is different in two isoforms of PPARγ. In contrast to PPARγ1, PPARγ2 has 28 additional amino acids at the amino-terminus. It is expressed in a more adipose-selective manner and about ten times more abundant than PPARγ1 (Tontonoz *et al.*, 1994). Furthermore, PPARγ2 appears to be more potent to induce adipogenesis (Rosen *et al.*, 2002). MAP kinase (specifically Erk1, also known as p44) mediated phosphorylation of AF1's Ser112 reduces ligand-binding affinity of the receptor and negatively regulates the transcriptional activity, resulting in the inhibition of its biological action (Adams *et al.*, 1997; Hu *et al.*, 1996). The phosphorylation was suggested to induce a conformational change of the unbound receptor, in which the amino-terminal AF1 inhibits, via intramolecular interactions, ligand binding to the LBD (Shao *et al.*, 1998). The AF1 was also shown to bind to PPARγ co-activator 2 (PGC-2), not affected by Ser112 phosphorylation, which itself exhibits pro adipogenic action (Castillo *et al.*, 1999). However, the exact mechanism of AF1 mediated regulation of PPARγ's adipogenic action remains still elusive and requires further investigation.

2.1.1.2 The ligand-binding domain (LBD)

The LBD has several regulator functions. It determines the subcellular localisation, initiates heterodimerisation with RXRα and regulates the transcriptional activation or repression in a ligand-dependant manner. The LBD consists of 13 α-helices and a small four-stranded β-sheet region (Figure 2.1-1D). The AF2, the receptors surface facilitating co-activator interaction, is formed by the LBD's helix 3-4 loop and the carbox-terminal ends of helix H11 and H12. The LBD has a phosphorylation site at Ser245 (Ser273 in PPARγ2) (Choi *et al.*, 2010) and possesses a cavity of about 1300 Å^3 in the lower portion of the LBD (Nolte *et al.*, 1998), located between helix H3 and the β-sheet region (Figure 2.1-1C-E).

Ligand binding to the LBD of the nuclear receptor leads to co-activator recruitment. Co-activators like the Steroid receptor co-activator-1 (SRC1) or the Peroxisome proliferator-activated receptor gamma co-activator-1 alpha (PGC-1α) possess an LXXLL motif (where L is leucine and X is any amino acid) that dock with their hydrophobic sides, formed by the leucine residues, into a hydrophobic groove on the surface of the receptor (Figure 2.1-2) (Heery *et al.*, 1997; Li *et al.*, 2008; Nolte *et al.*, 1998). A positively charged lysine (K301) and a negatively charged glutamic acid (E471) residue at each side of the hydrophobic groove contribute to the selectivity by clamping the LXXLL motif by hydrogen bonds with the motif's backbone at both, the amino- and carboxy-terminal ends (Darimont *et al.*, 1998; Nolte *et al.*, 1998; Shiau *et al.*, 1998). This kind of interactions of the hydrophobic groove with the co-activator motif leads to the name 'charge clamp'. In the unbound state, nuclear receptors are

frequently blocked by co-repressors like the silencing mediator co-repressor for retinoid and thyroid-hormone (SMRT) or the nuclear receptor co-repressor (N-CoR) (Chen and Evans, 1995; Horlein *et al.*, 1995) containing an LXX[I/H]IXXX[I/L] motif (Hu and Lazar, 1999; Nagy *et al.*, 1999; Perissi *et al.*, 1999). The ligand dependant distinction between co-repressor and co-activator was proposed to be caused by the different length of the motifs that have to be accommodated in the binding pocket. Ligand binding leads to a conformational change in which the amphipathic helix H12 is stabilised tightly packed against the LBD (see also section 2.1.1.3) and shortens the hydrophobic groove, such that it only suits the smaller co-activator motif (Perissi *et al.*, 1999).

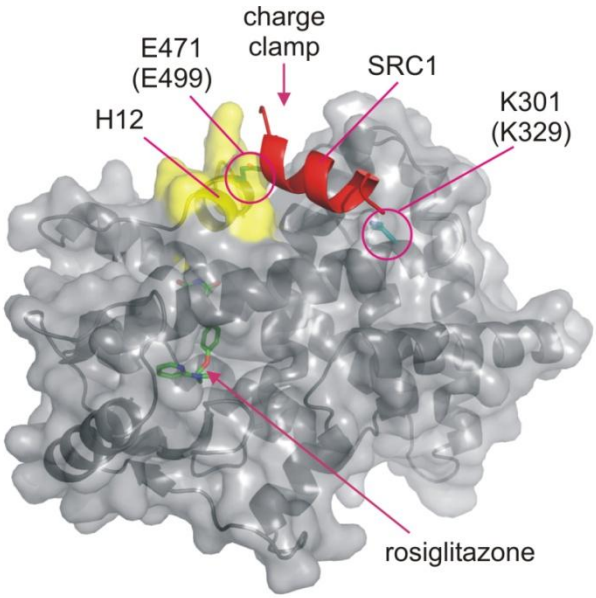

Figure 2.1-2: Complex of PPARγ's LBD with an SRC1 peptide. The SRC1 peptide (red) interacts with the charge clamp of the LBD (PDB 2PRG) in response to a ligand like rosiglitazone. The accessibility of the charge clamp is regulated by helix H12 (yellow) in a ligand dependant manner. Residues defining the charge clamp are indicated. Residues in parentheses correspond to residues in PPARγ2.

2.1.1.3 Ligand-mediated LBD stabilisation regulates PPARγ's transcriptional activity

Crystallographic temperature factors (B-factors) are linearly related to the mean square of an atom's displacement and indicate atomic flexibility. Crystal structures of unbound PPARγ have high temperature factors in their lower portion (Figure 2.1-1D), which thus seems to be less rigid than the upper portion. Helix H12 is the most mobile helix and the omega-loop, linking helix H2' and helix H3, is the most mobile loop (Nagy *et al.*, 1999; Nolte *et al.*, 1998). Both elements are located in the lower portion. Ligand binding to PPARγ reduces atomic thermal vibrations and stabilises the receptor. This stabilisation induces a conformational

change of helix H12, leading to a tight association of this helix to the LBD and results in a more compact and rigid structure (Choi *et al.*, 2010; Nagy and Schwabe, 2004). This affects, as described above, co-regulator recruitment. The ligand dependant degree of LBD stabilisation is believed to be directly proportional to the receptors transcriptional activity (Nagy and Schwabe, 2004) and may cause differential co-regulator recruitment (Choi *et al.*, 2010). Stabilisation can be achieved by direct contacts of the ligand to helix H12, as shown for full agonists like rosiglitazone (see Figure 2.1-4 and section 2.1.1.3.1), or by a helix H12 independant stabilisation, as shown for partial and intermediate agonists like the non-thiazolidinedione selective partial agonist (nTZDpa) and BVT.13 (see Figure 2.1-3 and section 2.1.1.3.2). However, both mechanisms of stabilisation result in a similar position of helix H12.

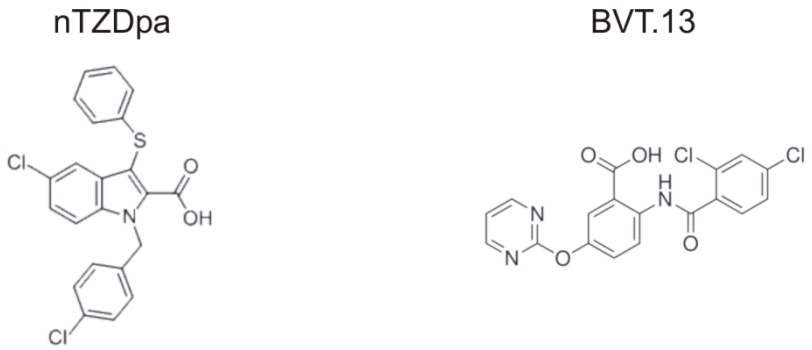

Figure 2.1-3: The PPARγ partial agonist nTZDpa and the intermediate agonist BVT.13.

2.1.1.3.1 The PPARγ full agonist rosiglitazone

The synthetic compound rosiglitazone belongs to the group of thiazolidinediones (TZDs) and is a potent, high-affinity PPARγ activator, which is used as insulin sensitizer to treat type II diabetes. The small, lipophilic rosiglitazone molecule targets adipose tissue where PPARγ is mainly expressed and stabilises preferentially helix H12 by direct contacts (Figure 2.1-4A). This kind of stabilisation leads to co-activator recruitment and is suggested to 'freeze' the receptor in a rigid conformation. Obesity-induced Cdk5-mediated phosphorylation of the receptor LBD's Ser245 is thereby inhibited (Choi *et al.*, 2010). The inhibition of the phosphorylation is apparently closely associated with the anti-diabetic effects of rosiglitazone, as it alters the expression of a subset of genes with regulatory functions in metabolism, like an enhanced transcription of the insulin-sensitising adipokine, adiponectin (Choi *et al.*, 2010). Rosiglitazone is currently used to treat type II diabetes, but causes adverse effects including weight gain, oedema, and congestive heart failure (Kahn and

McGraw, 2010; Lebovitz, 2002). These adverse effects rationalise research to explore new agonists for PPARγ which also induce insulin sensitivity but lack these adverse effects.

2.1.1.3.2 PPARγ partial and intermediate agonists

Full, intermediate, and partial agonists are defined by transcriptional activation potentials of >80%, 50-80%, or <50%, relative to the full agonist rosiglitazone. In contrast to full agonists, intermediate and partial agonists like BVT.13 and nTZDpa or MRL-24 (see Figure 2.1-4) do not stabilise helix H12 by direct contacts but rather globally stabilise the LBD by contacts to helix H3 and the β-sheet region of the binding pocket in the lower portion of the domain (Figure 2.1-4) (Bruning *et al.*, 2007). The mode of stabilisation of the receptor by partial agonists leads to a more effective inhibition of Ser245 phosphorylation and to a more specific and in general reduced gene activation profile, compared with the full agonist rosiglitazone. These effects were already described for the partial agonist MRL-24 (Choi *et al.*, 2010). Partial agonists, also known as selective PPARγ modulators (SPPARγMs) are able to separate adverse effects like adipogenesis from inducing insulin sensitivity in vivo (Rangwala and Lazar, 2002).

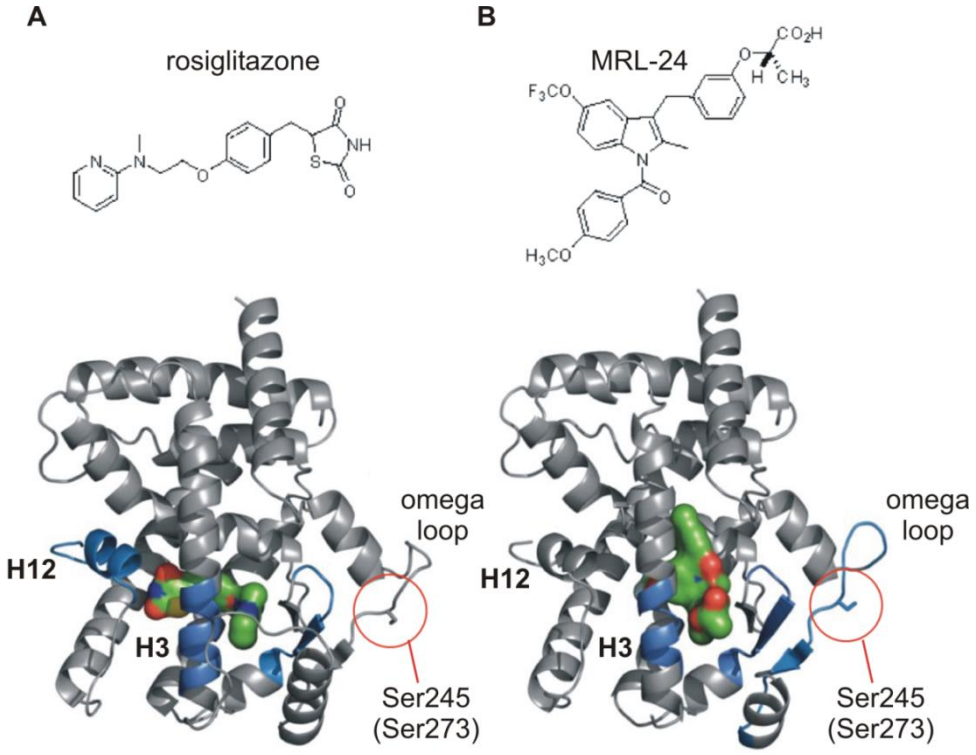

Figure 2.1-4: Ligand dependant stabilisation of the LBD. LBD stabilisation by the full PPARγ agonist rosiglitazone (A) and the partial agonist MRL-24 (B) according to the hydrogen/deuterium exchange data of Choi *et al.* 2010. The LBD's stabilised regions are coloured blue. Helix H12, H3, and the phosphorylation site are indicated. Note the differentially stabilisation of helix H12 and the omega loop, containing the phosphorylation site.

Co-activators like SRC-1 are known to bind to the charge clamp at helix H12 in response to ligands like the full agonist rosiglitazone. But, as there are no significant conformational differences observed between the LBD bound to full, intermediate or partial agonists, it was suggested that the differentially stabilisation of the receptor and in particular the diminished stabilisation of helix H12 by partial agonists is reasonable for the reduced and more specific gene activation profile. Furthermore, the phosphorylation of Ser245 on the opposite site of the domain may constitute a second ligand-gated co-regulator binding site near this residue (Choi *et al.*, 2010). The differentially regulation of this phosphorylation site by partial or intermediate agonists may thus also affect co-regulator recruitment. Agonists with the potential to induces insulin sensitivity with milder or even without the induction of adverse effects may pave the way to more specific PPARγ activation for the clinically use as therapeutic drugs to treat diseases like type II diabetes, while maintaining the patients quality of life.

2.1.1.4 Co-regulator recruitment

It was suggested that distinct ligands generate by differential stabilisation unique co-activator-binding surfaces on the LBD, resulting in a differential recruitment of co-regulators and thus a different transcriptional activation by PPARγ (Choi *et al.*, 2010; Kodera *et al.*, 2000; Wu *et al.*, 2003). Furthermore, sequences flanking the LXXLL motif of the co-activator also affect the recognition by the receptor (Klein *et al.*, 2005; Wu *et al.*, 2003). These suggestions are further supported by the findings that PGC-1α is able to act as a co-activator of PPARγ in the presence or absence of rosiglitazone (Li *et al.*, 2008; Wu *et al.*, 2003). While the rosiglitazone-mediated co-activation is independant of the charge clamp mutation E471A, the ligand-independant activation is abolished by this mutation. This indicates a unique, ligand-dependant structural feature of the LBD that associates with PGC-1α (Wu *et al.*, 2003). On the contrary, the rosiglitazone mediated activation is inhibited by the PGC-1α mutant S(-2)A, with the -2 position relative to the first leucine of the ^{144}LXXLL^{148} motif (Wu *et al.*, 2003). Thus both, the ligand but also the co-activator configurations are critical for co-regulator recognition and nuclear receptor activity.

2.1.2 A new class of high affinity PPARγ partial agonists

A new class of high affinity PPARγ partial agonists, the amorfrutins (see Figure 2.1-5), was identified by our cooperation partner, the group of Dr. Sascha Sauer, Max Planck Institute for Molecular Genetics, Berlin. They have screened a library of about 8000 pure natural products using an electrospray ionisation mass spectrometry based high-throughput

screening method and identified the amorfrutins as structurally new and affine PPARγ agonists. Amorfrutins are a group of natural products with the characteristics of SPPARγMs (Weidner *et al*., submitted). They are non-toxic ingredients of edible plant roots of liquorice, *Glycyrrhiza foetida*, and fruits of a related legume, A*morpha fruticosa*, also known as desert false indigo or bastard indigobush. The name 'amorfrutin' is derived from *Amorpha fruticosa*, in which the molecules were originally identified. Amorfrutins are selective and high affinity PPARγ agonists with dissociation constants for amorfrutin 1, 2, and B of about 20-300 nM, determined by time resolved FRET assays. The small, lipophilic amorfrutin molecules consist of a 2-hydroxy benzoic acid core structure with diverse phenyl and isoprenyl moieties. Amorfrutin 1 was shown to increase insulin sensitivity without the induction of adipogenesis and even lead to weight loss in vivo (Weidner *et al*., submitted). Amorfrutins have great potential for treatment or prophylaxis of type II diabetes and the metabolic syndrome in general.

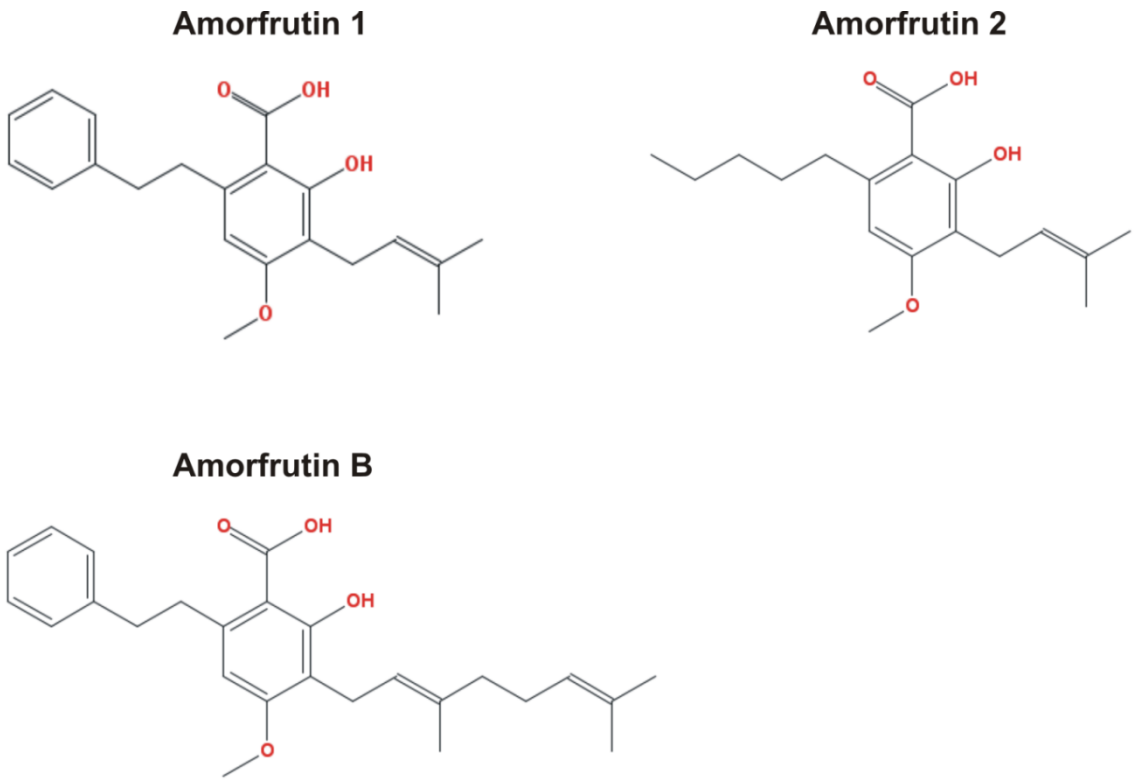

Figure 2.1-5: Structures of amorfrutin 1, 2, and B, used in this study.

2.2 Aim of this work

The aim of this work was to solve the crystal structures of amorfrutin 1, 2, and B in complex with the ligand-binding domain (LBD) of human PPARγ for structural characterisation of these novel natural products at the atomic level. Recent research regarding ligand specific selectivity, affinity, and transcriptional activation (Weidner *et al.*, submitted) led to the assumption that different structural features of amorfrutins stabilise the transcriptional receptor. Structural comparison, also with related complex structures, should help to identify these features, elucidate the mechanism of activation, and classify, whether amorfrutins are full, intermediate, or partial agonists. Furthermore, the study should explain how the structural data will fit into the current working model of PPARγ activation. The results of this study should contribute to the understanding of the regulatory mechanism of PPARγ in insulin sensitisation and fat metabolism.

2.3 Materials

If not stated otherwise, the materials used in this chapter are similar to the materials in chapter I.

2.3.1 Plasmid for hPPARγ LBD expression

A PCR fragment encoding amino acids 223-505 (AEIS...KDLY) of the LBD of hPPARγ (SwissProt PPARG_HUMAN P37231), cloned with NdeI and BamH1 into a pET-28a vector (Novagene), was used for expression, resulting in hPPARγ LBD fused to an amino-terminal His_7-tag and thrombin cleavage site, with the sequence GSHMAEI...KDLY upon cleavage.

2.3.2 Compounds

The compounds amorfrutin 1 (NP-003520), amorfrutin 2 (NP-003521), and amorfrutin B were synthesized by Frank C. Schröder (Boyce Thompson Institute and Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, USA) and obtained dissolved in DMSO at 100 mM. Higher concentrations of the amorfrutins couldn't be achieved due to their solubility limits.

2.4 Methods

If not stated otherwise, the methods used in this chapter are similar to the methods in chapter I.

2.4.1 Production and purification of PPARγ's LBD

The PPARγ LBD was expressed with a previously described expression plasmid kindly provided by Krister Bamberg (Cronet *et al.*, 2001). Transformed BL21(DE3) cells were induced with 0.1 mM IPTG at 18°C for 20 h. Harvested cells were disrupted with a high pressure cell disrupter (Constant Systems) in 20 mM Tris-HCl, 150 mM NaCl, 10% (v/v) glycerol, 1 mM tris(2-carboxyethyl)phosphine HCl (TCEP, Gold Biotechnology), 10 mM imidazole, pH 8.0 in the presence of protease inhibitors (Complete EDTA-free, Roche) and centrifuged (35,000xg, 45 min). The supernatant was loaded on a 5 ml HisTrap HP column (GE Healthcare) and eluted with an imidazole gradient (10-300 mM imidazole) over 20

column volumes. The protein was diluted to a final concentration of 20 mM NaCl, immediately loaded onto a MonoQ HR 10/10 column (GE Healthcare) and eluted with a NaCl gradient (20-500 mM NaCl), followed by incubation with thrombin protease (1 U/mg) at 4°C for 20 h to cleave off the His-tag. His-tag peptides and uncleaved material were removed by rechromatography with Ni-NTA agarose (Qiagen), followed by gel filtration (HR 10/30 Superdex 75 prep grade column, GE Healthcare) in 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM DTT, and 2 mM EDTA, pH 8.0.

2.4.2 Analytical methods

Photometric quantification of protein concentrations, SDS-PAGE, DLS, and MS were carried out in accord with the description in chapter I (see section 1.4.3).

2.4.2.1 Analytical gel permeation chromatography

Analytical gel permeation chromatography was performed in 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM DTT, and 2 mM EDTA, pH 8.0 using a HiLoad 16/60 Superdex 75 column (GE Healthcare). 1 ml buffer containing a protein mixture of 20 mg/ml ribonuclease A (14.6 kDa), 20 mg/ml ovalbumin (46.7 kDa), and 10 mg/ml thyroglobulin (699 kDa) was applied to the column and used as molecular weight standard. The gel-phase distribution coefficient (K_{av}) was calculated according to Equation 2.4-1, where V_0 is column void volume, V_e is elution volume, and V_c is geometric column volume, calculated from column dimensions. V_0 was approximated from the elution volume of thyroglobulin. K_{av} was plotted versus the logarithm of the molecular weight of the different standard proteins. The equation of the linear fit of the relationship between K_{av} and the different molecular weights of the standard proteins was calculated and used to estimate the molecular weight of LBD.

$$K_{av} = \frac{V_e - V_0}{V_c - V_o}$$ **Equation 2.4-1**

2.4.2.2 AF4 and DLS

Protein samples of the LBD were analysed by Asymmetric Flow Field-Flow Fractionation (AF4), coupled to a dynamic light scattering (DLS) device, in-house by Dr. Thorsten Lührs.

2.4.3 Crystallisation by vapour diffusion

2.4.3.1 Sitting-drop vapour diffusion

The sitting-drop vapour diffusion method was used for initial screening in 96-well format as described in chapter I (see section 1.4.4.1).

2.4.3.2 Hanging-drop vapour diffusion

The protein was concentrated to 11 mg/ml and crystallised using hanging drop-vapour phase diffusion by mixing equal volumes of concentrated protein and reservoir solution (0.8 M tri-sodium citrate and 0.1 M imidazole, pH 8.0). Crystals appear after 2-5 days at 19°C.

2.4.3.3 Crystal soaking

Soaking crystals with the different amorfrutins was carried out using crystals grown under conditions described in section 2.4.3.2. The crystals were transferred into soaking buffer (10 mM Tris-HCl, 50 mM NaCl, 0.25 mM DTT, 1 mM EDTA, 0.8 M M tri-sodium citrate, 0.1 M imidazole, final pH 8.0) with a stepwise increase in reagent concentration (0.5 mM, 1 mM, and 1.5 mM ligand and a maximum of 1.5% DMSO) with varying incubation times.

2.4.3.4 Co-crystallisation

The complexes of the PPARγ LBD bound to the different amorfrutins were prepared by adding a 10-fold molar excess of each amorfrutin to the purified LBD at a final concentration of 1 mg/ml. DMSO and unbound ligand were removed by using a HR 10/30 Superdex 75 prep grade column (GE Healthcare) in 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM DTT, and 2 mM EDTA, pH 8.0. The complexes were crystallised using the hanging-drop vapour diffusion method according to section 2.4.3.2, but had to be optimised by microseeding (see section 1.4.4.3).

2.4.4 Data collection, structure determination and refinement

To prevent or minimise crystal damage by radiation, the crystals were transferred into a mild hypertonic cryo-protectant solution (0.84 M tri-sodium citrate, 25% (v/v) glycerol, and 0.1 M imidazole, pH 8.0) and immediately thereafter flash frozen in liquid nitrogen before data collection at 100 K. The data sets were collected in-house on a Saturn 944+ detector

(Rigaku) using Cu Kα X-rays from a Rigaku MicroMax-007 HF rotating-anode X-ray generator with a VariMax Optic (Rigaku). Images were indexed and processed with XDS (Kabsch, 1993) and the structures were solved by molecular replacement using CCP4 MOLREP (Vagin and Teplyakov, 1997) with PDB entry 1PRG (Nolte *et al.*, 1998) as the search model. REFMAC5 (Murshudov *et al.*, 1997) and Phenix.refine (Adams *et al.*, 2010) were used for refinement. TLS (translation/libration/screw) parameters (Painter and Merritt, 2006) were included to allow domain movements to be refined and Coot (Emsley and Cowtan, 2004) was used for manual model building. Figures were prepared with PyMOL. Root mean square deviations (r.m.s.d.) between common Cα-positions were calculated with ProFit using the McLachlan algorithm (McLachlan, 1982). The validation of the structure was performed using MolProbity (Chen *et al.*, 2010; Davis *et al.*, 2007).

2.4.5 Bioinformatics

Structural analysis and figure preparations were carried out according to chapter I (see section 1.4.6).

2.4.5.1 Sequence analysis

Protein sequences were aligned with ClustalW2 (Larkin *et al.*, 2007) and displayed with ESPript 2.2 (Gouet *et al.*, 2003).

2.5 Results

2.5.1 Production and purification of the PPARγ ligand-binding domain (LBD)

The PPARγ LBD was expressed and purified according to section 2.4.1, using IMAC as the initial purification step. The LBD was eluted from the matrix using a gradient up to 300 mM imidazole. The protein was already completely eluted after 130 mM imidazole (Figure 2.5-1). The yield of IMAC purified PPARγ LBD was about 70 mg/L culture.

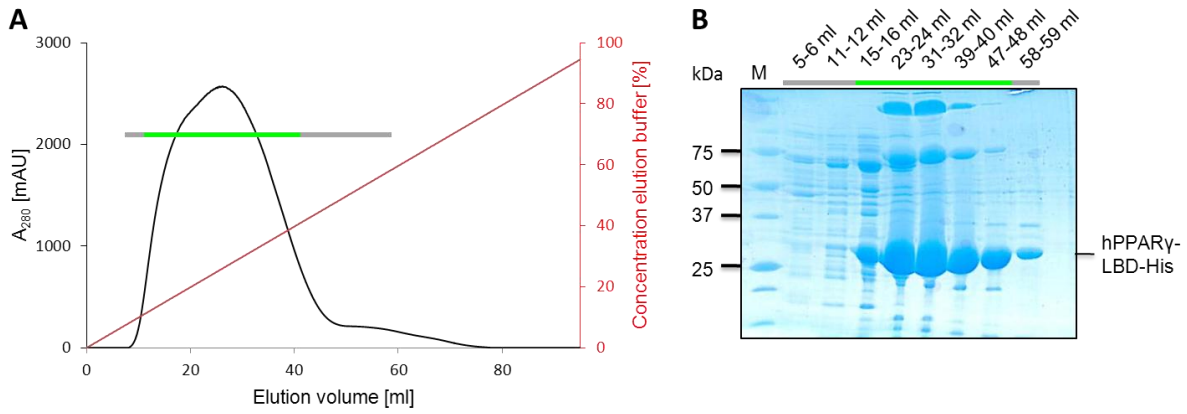

Figure 2.5-1: IMAC purification of the PPARγ LBD. Bacterial lysate containing the LBD was loaded on a 5 ml HisTrap HP column and eluted with a gradient of 10 to 300 mM imidazole (A). The grey and green line depicts the fractions analysed by SDS-PAGE and Coomassie Blue staining (B). The fractions marked green were pooled for further purification. M: protein standard.

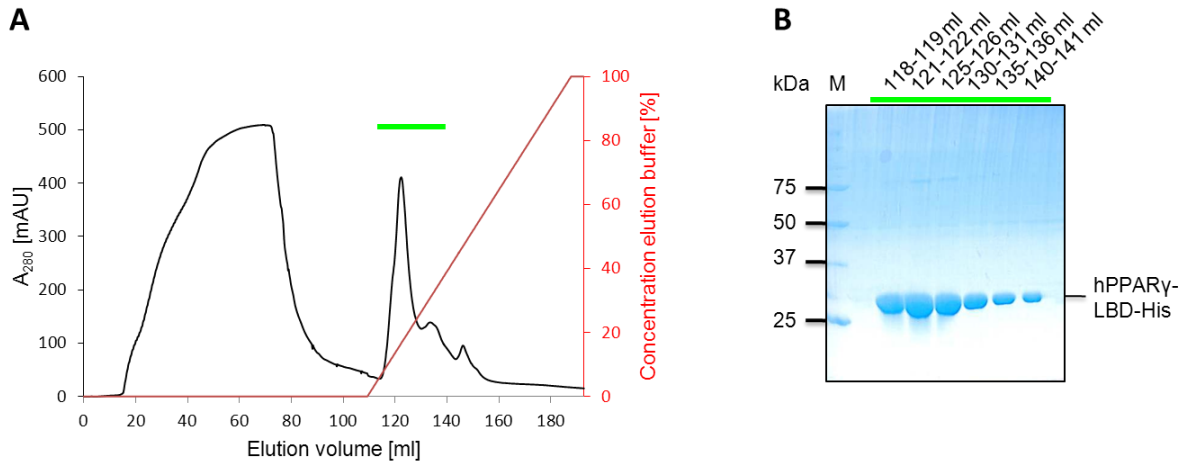

Figure 2.5-2: Anion exchange purification. The PPARγ LBD purified by IMAC was loaded on an anion exchange MonoQ HR 10/10 column and eluted with a gradient of 20 to 500 mM NaCl (A). The green line marks the fractions analysed by SDS-PAGE and Coomassie Blue staining (B). The marked fractions were pooled for further processing. M: protein standard.

The LBD was further purified and concentrated by anion exchange chromatography. It was eluted early in a gradient of 20 mM to 500 mM NaCl (Figure 2.5-2). The ion exchange chromatography leads to a high purity of the protein, which was subsequently digested using thrombin protease to cleave off the His_7-tag (Figure 2.5-3). IMAC Rechromatography with Ni-NTA superflow was used to remove the His-tag peptide and uncleaved material. Gel filtration was used for final purification and buffer exchange (Figure 2.5-4).

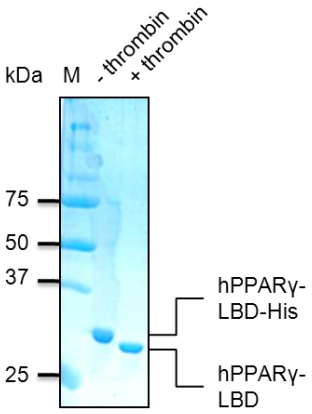

Figure 2.5-3: SDS-PAGE verifying successful removal of the approximately 3 kDa His_7-tag. M: protein standard.

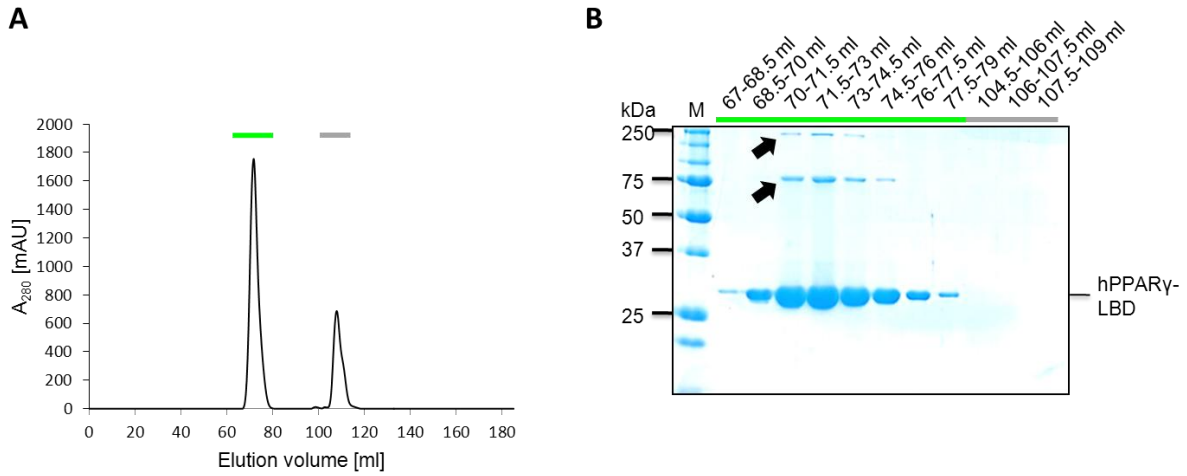

Figure 2.5-4: Gel filtration. The LBD of PPARγ elutes at 71 ml from a Superdex 75 HR 16/60 column (A). The grey and green lines depict the fractions analysed by SDS-PAGE (B), whereas the fractions marked green were pooled. Arrows indicated artefacts of the gel electrophoresis. M: protein standard.

SDS-PAGE of the gel filtration fractions resulted in two additional bands at 75 kDa and 250 kDa next to the expected 32.7 kDa LBD band. Mass spectrometry, analytical gel filtration as well as AF4, coupled to DLS, were carried out to identify these additional bands. Mass

spectrometry, performed in-house by Dr. Manfred Nimtz and Undine Felgenträger, showed that all three bands consisted of hPPARγ LBD. Furthermore, analytical gel filtration and AF4-DLS both revealed a monodisperse protein solution (Figure 2.5-4A and Figure 2.5-5). Analytical gel filtration, using thyroglobulin to approximate the void volume of the column and ovalbumin as well as ribonuclease A as molecular standards, resulted in a single peak with a calculated molecular weight of about 40 kDa (Figure 2.5-4A and Figure 2.5-5A,B). The AF4-DLS, as the much more precise method, was carried out in-house by Dr. Thorsten Lührs and revealed unambiguously a monodisperse protein solution and the exact molecular weight of 32.6 kDa for the LBD (Figure 2.5-5C). These data verify the desired purity and monodispersity of the monomeric protein in solution and confirm an artificial oligomerisation of the protein during the SDS-PAGE, due to the formation of stable multimers under the condition of the SDS-sample-buffer used for the SDS-PAGE.

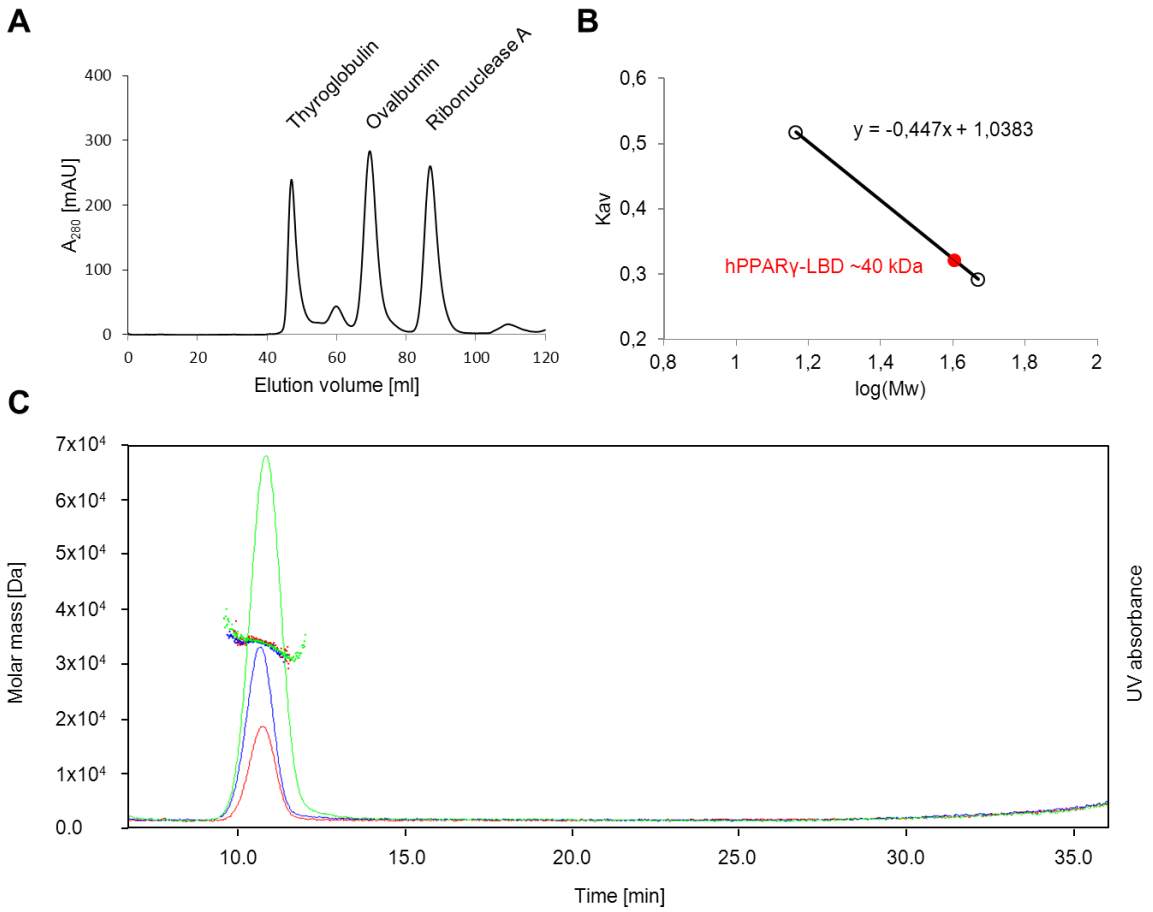

Figure 2.5-5: Monodispersity of the purified LBD of PPARγ. Elution profile of the analytical gel filtration (A) and the derived standard line (B). The standard line was calculated from ovalbumin and ribonuclease A elution volumes. Thyroglobulin was used to approximate the column's void volume. Using the standard line, a molecular weight of about 40 kDa was derived for the LBD. (C) AF4-DLS profile of the LBD. Shown are UV absorbance (continuous lines) and the molar mass (dashed lines) as a function of time. The analyses were performed using different protein amounts: 10 µg (red lines), 20 µg (blue lines), and 50 µg (green lines), and reveal the exact molecular mass of the LBD (32.6 kDa).

2.5.2 Crystallisation of PPARγ's LBD with different amorfrutins

Protein crystals of the LBD in complex with the different amorfrutins were obtained by crystal soaking (section 2.5.2.1) and co-crystallisation (section 2.5.2.2), respectively.

2.5.2.1 Crystal soaking

Crystal soaking is fast and more convenient for obtaining crystal structures of one protein with different ligands, as it requires only one crystallisation condition. Thus, crystallisation conditions for ligand-free PPARγ LBD were screened based on the previously published crystallisation conditions of this protein (Nolte *et al.*, 1998) using the hanging-drop vapour phase diffusion method (section 1.4.4.2), but also using the sitting-drop vapour phase diffusion method in 96-well format and the JCSG screens (section 1.4.4.1). Protein crystals appeared only with sodium citrate as the precipitant. Crystallisation conditions were optimised to 0.8 M tri-sodium citrate and 0.1 M imidazole, pH 8.0 (section 2.4.3.2). Protein crystals appeared already after two days at 19°C. Within two weeks, some of the crystals grew up to large monoclinic crystals with edge lengths of about 400 µm (Figure 2.5-6). The protein crystals diffracted X-rays up to 2.0 Å resolution.

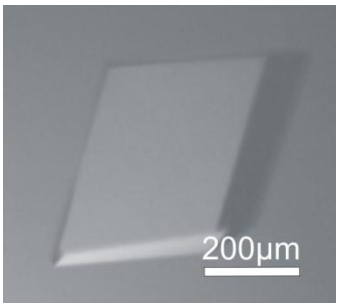

Figure 2.5-6: Crystal of the ligand-free LBD of PPARγ. The scale bar corresponds to 200 µm.

The amorfrutins were obtained dissolved in DMSO at a maximal concentration of 100 mM, due to their solubility limits. To avoid DMSO mediated crystal damage and to determine the maximal amorfrutin concentrations that could be used during the soaking, the effect of the organic solvent on crystal integrity was tested prior to the soaking of the compounds.

The crystals were transferred into soaking buffer with different DMSO concentrations. After 24 h incubation, damage of the crystals became visible at DMSO concentrations above 1.5%, like rounded edges or crystal cracks (Figure 2.5-7A). These effects did not exacerbate even after three days of incubation. Thus, a maximum of 1.5% DMSO and hence a maximum

concentration of the amorfrutins of 1.5 mM was used during the stepwise soaking of the LBD crystals. However, crystal soaking with the different amorfrutins resulted in crystal damage even at 1.5% DMSO (Figure 2.5-7B) and a decreased diffraction from up to 2.0 Å to about 2.5 Å, even at lower concentrations and varying incubation times. The resulting electron densities turned out to be too badly defined to identify the precise position and molecular interactions of the amorfrutins in the LBD.

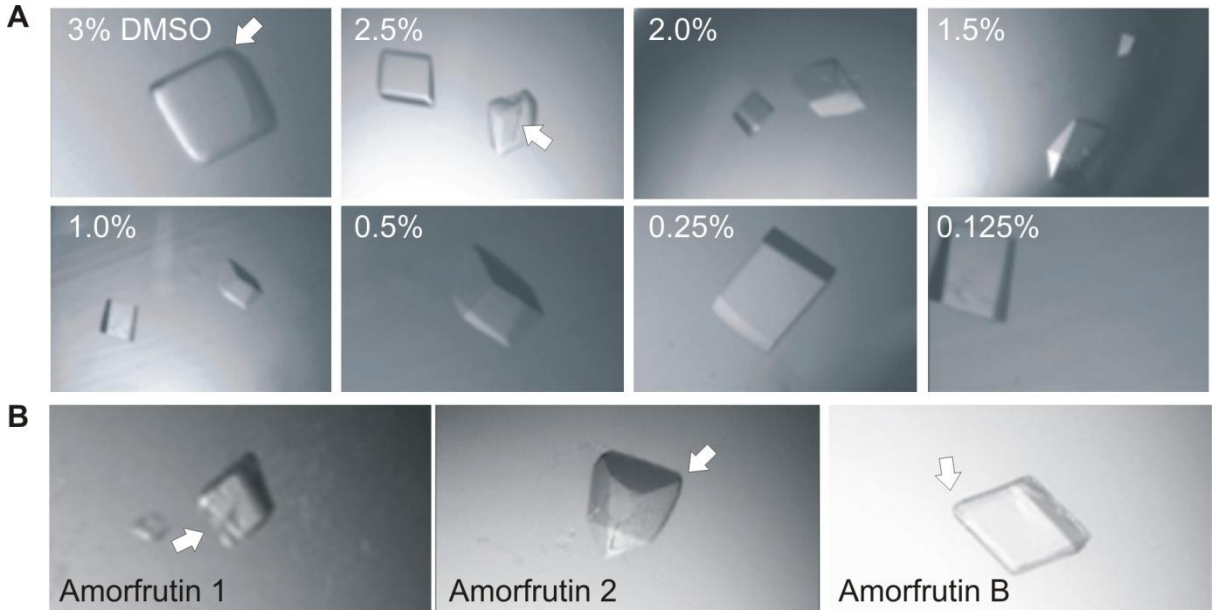

Figure 2.5-7: The effects of soaking on crystal integrity. (A) DMSO lead to rounded edges and crystal cracks at concentrations above 1.5% and an incubation time of 24 h. (B) The amorfrutins, dissolved in DMSO, lead to affected crystal integrity even at DMSO concentrations below 1.5% and varying incubation times. Crystal damages are indicated by arrows.

2.5.2.2 Co-crystallisation

In comparison to crystal soaking, much better data were obtained by co-crystallisation of amorfrutins with PPARγ. The complexes of the LBD bound to different amorfrutins were formed according to section 2.4.3.4. The crystallisation conditions were similar to the crystallisation of the ligand free protein, but resulted in adnate, distorted, and cracked crystals. However, the microseeding technique was used to optimise crystal quality, leading to diffraction up to 2.0 Å and a well-defined electron density map that allowed the precise definition of the amorfrutins in the ligand binding pocket of PPARγ and the description of all molecular interactions.

2.5.3 Structural characterisation of different amorfrutins in complex with the LBD of PPARγ

The crystal structures of the LBD in complex with three different amorfrutins were solved by molecular replacement and were all refined at a resolution of 2.0 Å (Table 2.5-1). The electron density maps reveal clearly the canonical three layer α-helical sandwich composed of 13 α-helices and a small, four-stranded β-sheet (Figure 2.5-8). The structures are in accord with previously published PPARγ LBD structures (Bruning *et al.*, 2007; Cronet *et al.*, 2001; Nolte *et al.*, 1998) with a calculated r.m.s. deviation of 0.87 ± 0.09 Å to Protein Data Bank (PDB) entry 1PRG (258 common Cα-positions of chain A aligned).

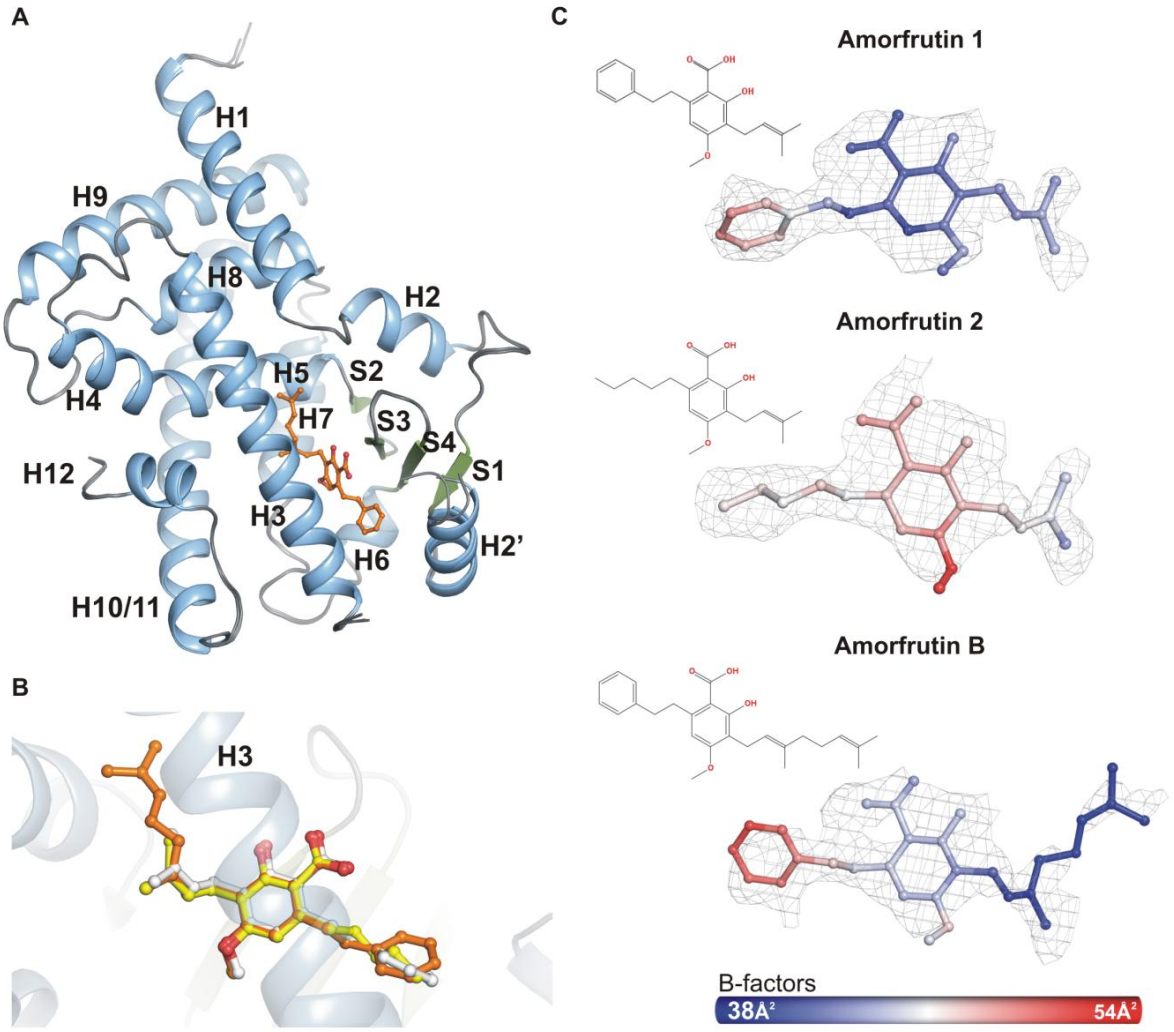

Figure 2.5-8: Structural comparison of the three amorfrutins bound to the LBD of PPARγ. (A) The aligned overall structures of the LBDs bound to the three amorfrutins are almost indistinguishable. Only amorfrutin B (orange sticks) is shown. The amorfrutins bind between the lower portion of helix H3 and the β-sheet region. Helices (blue) and β-strands (green) are labelled H1 to H12 and S1 to S4, respectively. (B) Superimposition of the three amorfrutins in the binding pocket, generated by alignment of the three respective protein structures. Amorfrutin 1 (white), 2 (yellow), and B (orange) are shown as sticks. Only the amorfrutin B-bound domain is displayed as ribbons. (C) $2F_O$-F_C electron density omit map calculated around amorfrutin 1, 2, and B, contoured at 1.0 σ. The amorfrutins dynamic is illustrated by plotting B-factors on the structures.

Table 2.5-1: Summary of crystallographic analysis.

Data collection	**amorfrutin 1**	**amorfrutin 2**	**amorfrutin B**
Space group	C121	C121	C121
Cell dimensions a, b, c (Å)	92.81, 61.10, 118.22	92.29, 60.97, 117.9	92.14, 60.86, 117.6
Unique angle β (°)	102.78	102.65	102.58
X-ray source	Rigaku MicroMax-007 HF		
Wavelength (Å)	1.5419		
Resolution range (Å)	27-2.0	27-2.0	29-2.0
Last shell (Å)	2.15-2.00	2.10-2.00	2.10-2.00
R_{merge} *(%)*	10.7 (64.4)	10.9 (56.3)	8.7 (45.4)
Observations	293322 (51030)	149072 (18878)	148018 (18473)
Unique reflections	43633 (8488)	43005 (5857)	42829 (5697)
Mean (I)/s.d. (I)	15.8 (3.1)	13.71 (2.82)	17.68 (3.41)
Completeness	99.5 (99.8)	99.0 (99.5)	99.2 (97.7)
Multiplicity	6.7 (6.0)	3.4(3.2)	3.5(3.2)
Structure refinement			
Resolution range (Å)	27-2.0	27-2.0	29-2.0
R_{work} (%)	20.4	20.89	20.3
R_{free} (%)	25.3	25.9	23.3
Total number of			
Non-hydrogen atoms	4554	4671	4670
Protein atoms	4299	4380	4339
Ligand atoms	75	44	60
Water molecules	180	247	271
r.m.s.d.			
Bond length (Å)	0.011	0.010	0.010
Bond angle (°)	1.240	1.219	1.186
Main chain B-factors (Å^2)	39.1	37.9	39.0
Side chain B-factors (Å^2)	42.2	41.3	42.3
Average B-factor protein atoms (Å^2)	40.7	39.7	40.7
Average B-factor ligand atoms (Å^2)	43.1	47.6	43.6
Average B-factor solvent atoms (Å^2)	44.3	44.9	44.3
Ramachandran statistics			
Most favoured regions (%)	98.8	97.3	98.3
Allowed regions (%)	1.2	2.7	1.7
Disallowed regions (%)	0	0	0

Values in parentheses are for the highest resolution shell.

All three complex structures had an asymmetric unit containing two amorfrutin molecules bound to one LBD homodimer of almost identical monomers. The structures of the two monomers, denoted chain A and B, differ due to crystal contacts. The two different conformations of the LBD chains, termed active and inactive, are commonly observed in PPARγ LBD structures (Cronet *et al.*, 2001; Nolte *et al.*, 1998). The inactive conformation (chain B) is caused by helix H12 of chain B forming crystal contacts with a neighbouring molecule that probably prevent it from moving into the active conformation like chain A (see also section 2.5.4). The active conformation of chain A is similar to the complexed receptor with the full agonist rosiglitazone and the co-activator SRC1 (Nolte *et al.*, 1998) and is commonly regarded as a suitable model for PPARγ LBD activation. The following sections will therefore refer only to the complex in the active conformation.

The electron density maps of each amorfrutin is well defined in the complex structures (Figure 2.5-8C) and clearly reveal their position in the cavity of the LBD. Corresponding to their related structures, the three amorfrutins are bound with almost identical localisation and orientation (Figure 2.5-8A and B). Whereas full agonists like rosiglitazone or MRL-20 stabilise especially helix H12 of the LBD (Bruning *et al.*, 2007; Nolte *et al.*, 1998), the amorfrutins stabilise the receptor between helix H3 and the β-sheet region, close to the ligand entry site. This recognition of the amorfrutins by PPARγ is similar to the partial agonists nTZDpa and MRL-24, and also to the intermediate agonist BVT.13 (Bruning *et al.*, 2007). These ligands all stabilise helix H3 and the β-sheet with hydrogen bonds to Ser342 and Arg288 and form similar extensive van der Waals contacts, most prominently to Ile341 of the β-sheet and Cys285 of helix H3 (Figure 2.5-9, Figure 2.5-10 and Figure 2.5-11A). The mechanism of stabilisation and the low transcriptional activation (<50%) compared to the full agonist rosiglitazone (Weidner *et al.*, submitted) (Table 2.5-2), imply that these amorfrutins belong to the group of partial agonists.

Despite their similar structures, the amorfrutins interact differently with the LBD. Arg288 of the amorfrutin 1-bound PPARγ adopts two alternative conformations (Figure 2.5-9A), which are also observed in nTZDpa-bound PPARγ, but not in the unbound structure (PDB entry 1PRG), where Arg288 is hydrogen bonded to Ser289 (Figure 2.5-10). In the amorfrutin 1 complex, Arg288 either interacts with Glu343 of the β-sheet and the ligand's carboxyl group, or with Glu295 of helix H3 via hydrogen bonds. Arg288 and Ser342 thus form a bond network upon binding of amorfrutin 1 that, supported by extensive van der Waals contacts of the phenyl and isoprenoid moieties, stabilises helix H3 and the β-sheet and explains the high binding affinity of 236 nM (Figure 2.5-9A and Table 2.5-2).

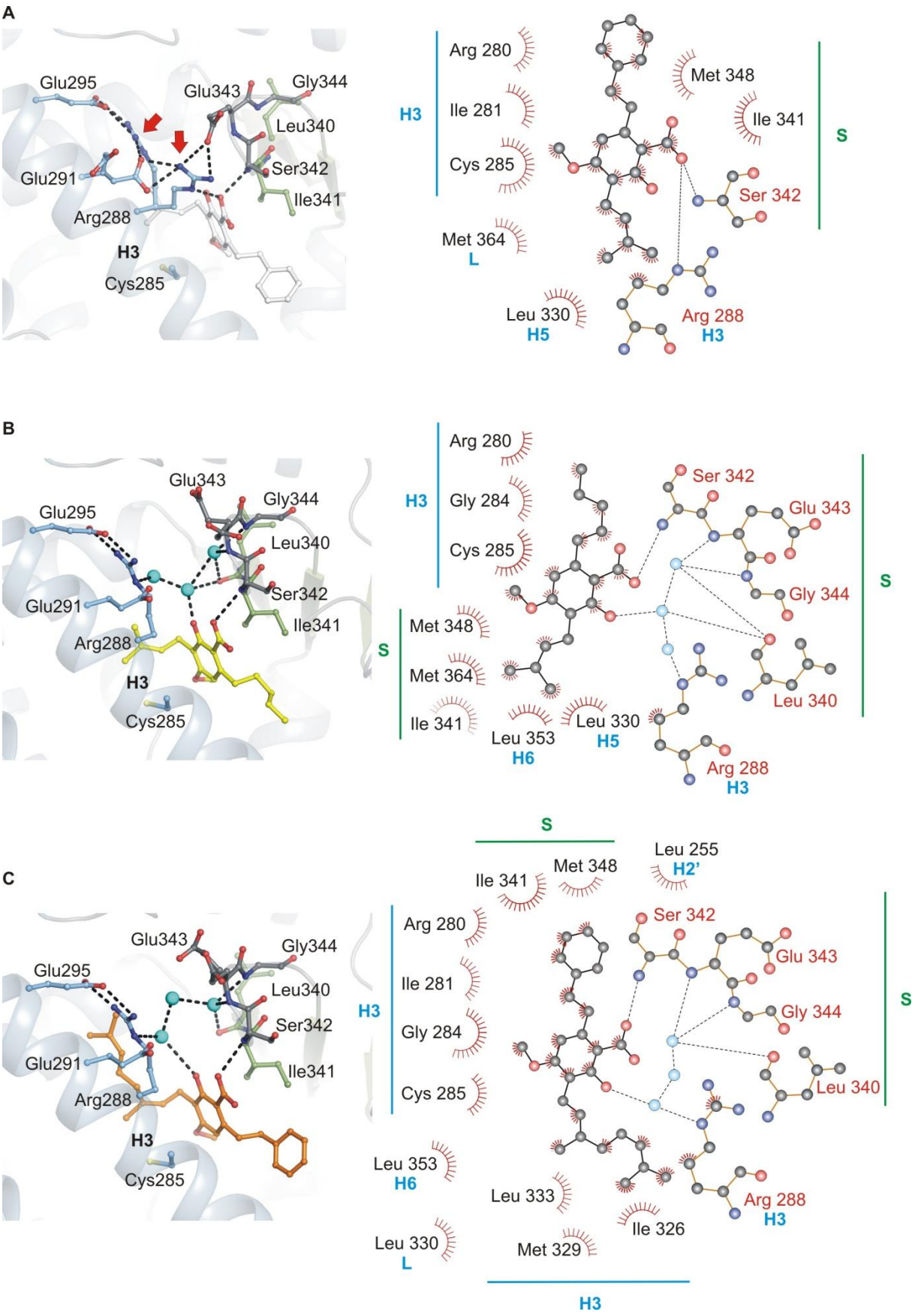

Figure 2.5-9: Structural details of amorfrutins bound to the LBD of PPARγ. (A) Amorfrutin 1 exerts a hydrogen-bonding network directly to Ser342 of the β-sheet region and Arg288 of H3. Note the alternative conformations of Arg288 (red arrows), stabilised by Glu291 and Glu343. (B, C) Amorfrutin 2 and B coordinate water molecules to form hydrogen-bonding networks between Arg288

of H3 and the β-sheet region, but also directly interact with Ser342. Additional hydrophobic interactions of amorfrutin B's longer geranyl side chain to Arg288 are the likely cause of this ligand's higher binding affinity and also prevent the formation of an alternative conformation of Arg288, as observed in the complex with amorfrutin 1. In the schematic diagrams of atomic interactions on the right, calculated using LIGPLOT (Wallace *et al.*, 1995), the ligands and the residues of the binding pocket are drawn with black and orange bonds, respectively. Hydrogen bonds are depicted by dashed lines and van der Waals contacts are indicated by spoked arcs and atoms with spokes. Oxygen atoms are coloured red, nitrogen atoms blue and carbon atoms black. S: sheet, H: helix, L: loop.

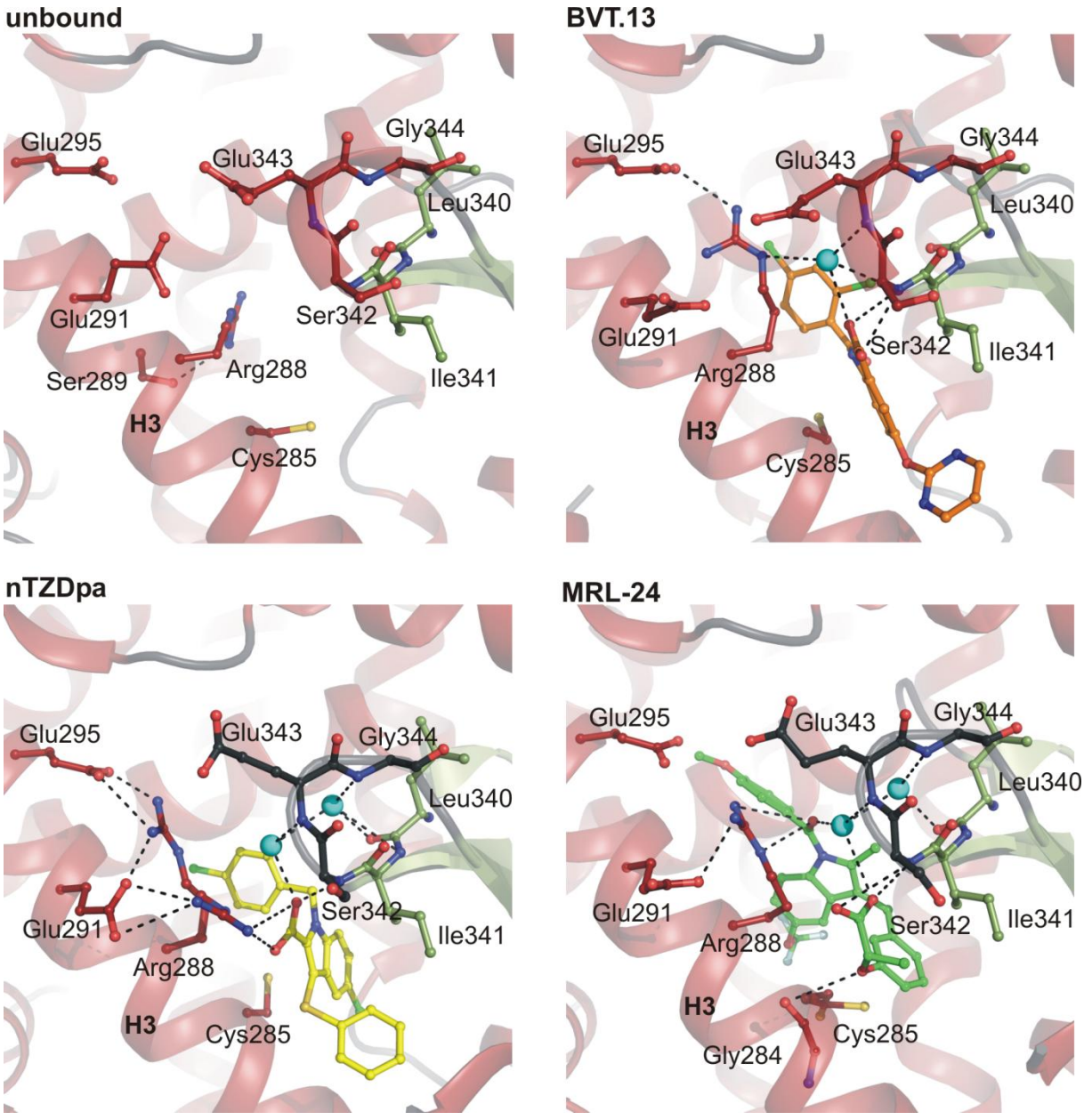

Figure 2.5-10: Comparison of previously published PPARγ structures. Shown are the unbound LBD of PPARγ (PDB 1PRG) with Arg288 hydrogen bonded to Ser289 (Nolte *et al.*, 1998), the LBD in complex with the intermediate agonist BVT.13 (PDB 2Q6S), and the LBD in complex with the partial agonists nTZDpa (PDB 2Q5S) and MRL-24 (PDB 2Q5P) (Bruning *et al.*, 2007).

Amorfrutin 2 and B also form direct hydrogen bonds to Ser342 of the β-sheet but fail to form a salt bridge to Arg288 and do not trigger the alternative conformation of this residue. The lack of this interaction results in a lower degree of stabilisation of the 2-hydroxy benzoic acid core of amorfrutin 2 and B, depicted by increased B-factors (see Figure 2.5-8). However, both ligands use the same residues of the LBD to form a water network that stabilises Arg288 of helix H3 (Figure 2.5-9B, C and Figure 2.5-11A), similar to the structure with BVT.13 of Bruning *et al.* (2007) (Figure 2.5-10). Amorfrutin 2 has a binding constant of 287 nM, similar to amorfrutin 1 with 236 nM (Table 2.5-2), whereas binding of amorfrutin B is markedly stronger with a binding constant of 19 nM.

Amorfrutin B's long meta-geranyl side chain at the 2-hydroxy benzoic acid core leads to additional hydrophobic contacts especially to Arg288 of helix H3 and to helix H4-5, compared with amorfrutin 1 and 2 (Figure 2.5-9 and Figure 2.5-11A). These additional contacts of amorfrutin B lead to reduced B-factor values of the meta-geranyl side chain (see Figure 2.5-8) and likely prevent that Arg288 adopts the alternative conformation, required for direct contacts to the ligand's carboxyl group. The additional hydrophobic contacts also explain the almost 12-fold higher binding affinity of amorfrutin B compared to amorfrutin 1 (Table 2.5-2).

Table 2.5-2: Affinity constants (Ki) and transcriptional activation relative to rosiglitazone according to ([1]) Weidner *et al.* (submitted) and ([2]) Bruning *et al.*, 2007.

	PPARα Ki	PPARδ Ki	PPARγ Ki	Relative transcriptional activation
Amorfrutin 1	27 µM[1]	27 µM[1]	0.236 µM[1]	39%[1]
Amorfrutin 2	25 µM[1]	17 µM[1]	0.287 µM[1]	30%[1]
Amorfrutin B	2.6 µM[1]	1.8 µM[1]	0.019 µM[1]	n.a.
nTZDpa	n.a.	n.a.	0.029 µM[1]	>50%[2]
MRL-24	n.a.	n.a.	n.a.	>50%[2]
BVT.13	n.a.	n.a.	n.a.	50-80%[2]
Rosiglitazone	7 nM[1]	n.a.	7 nM[1]	100%[1]

n.a., not available

Amorfrutins exhibit high selectivity for PPARγ, as shown for the three amorfrutins with a selectivity factor of 60 to 140 in comparison to the other PPAR subtypes. Even if the conservation of the residues of the lower portion of helix H3 and the β-sheet is quite low (Figure 2.5-11) most of the interactions to the amorfrutins should also be fulfilled by the residues of PPARα or PPARδ. However, the guanidinium group of the unconserved Arg288 (Figure 2.5-11) seems to play a key role in amorfrutin binding and this feature cannot be satisfied by any other residue. Thus the missing conservation of Arg288 is most likely the reason of the high selectivity of the three amorfrutins to the PPARγ subtype.

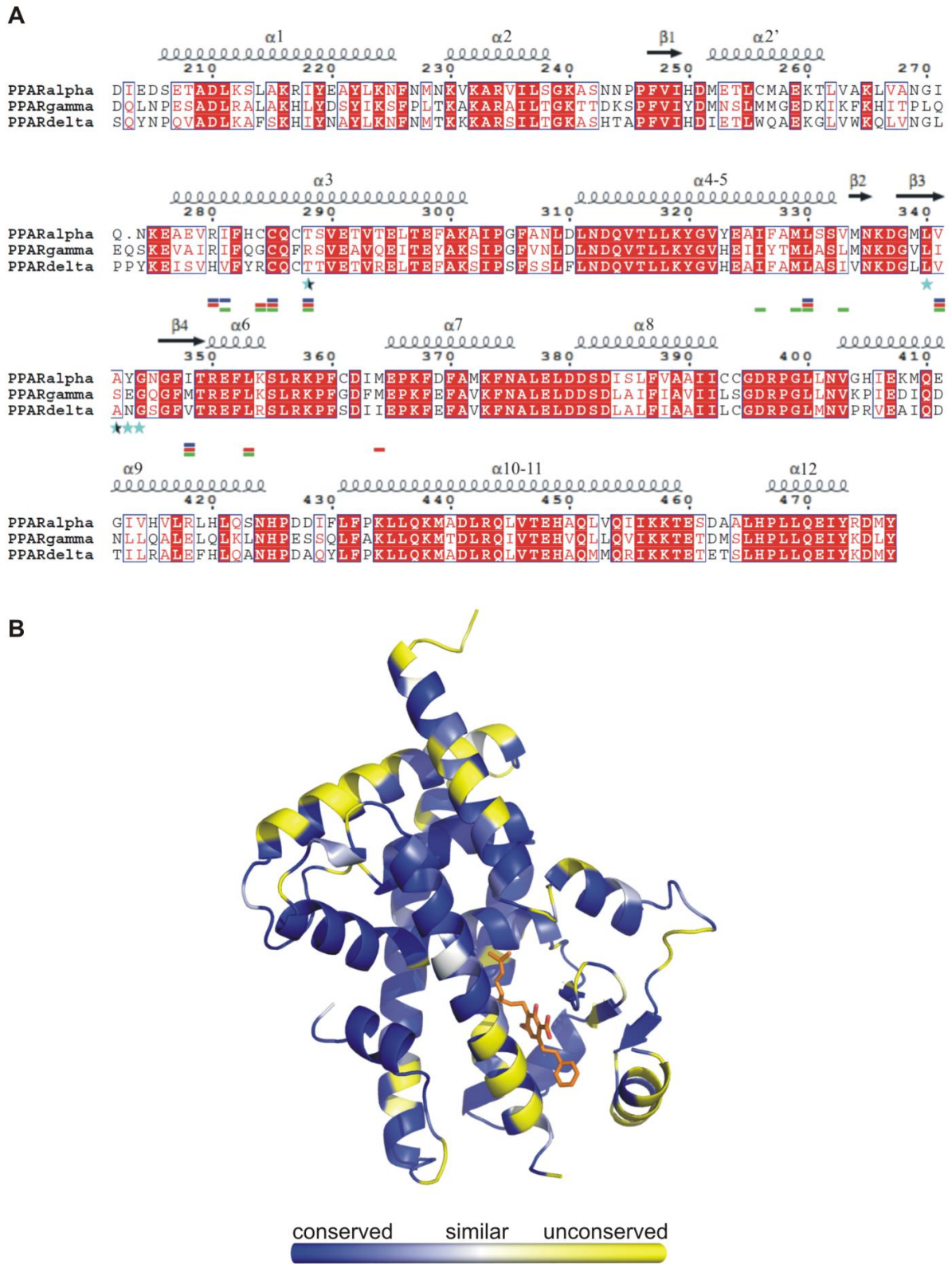

Figure 2.5-11: PPAR sequence conservation. (A) Sequence alignment of the three human PPARs, PPARα, PPARγ, and PPARδ (SwissProt Q07869, P37231, Q03181). The sequence identity to PPARγ is 69% and 72% for PPARα and PPARδ, respectively. Red and white boxes highlight identical and similar residues, respectively. Hydrophobic contacts of amorfrutin 1, 2, and B to PPARγ are indicated by blue, red, and green lines, respectively, beneath the sequences. Hydrogen bonds to amorfrutin 1 are indicated by black asterisks; hydrogen bonds to amorfrutin 2 and B are indicated by cyan asterisks. Ruler is given according to PPARγ PDB entry 1PRG. (B) Sequence conservation between the three different PPARs. Blue areas indicate identity, white areas similarity, and yellow areas represent differing residues. Similarity score is projected on the complex structure of the LBD with amorfrutin 1.

2.5.4 Crystallographic contacts of the inactive conformation lead to an artificial orientation and localisation of the amorfrutins

The orientation and localisation of the amorfrutin complex structures seem to be unaffected by crystallographic contacts in the active conformation. The inactive conformation (chain B), however, is obviously altered due to several contacts to a crystallographic related molecule. Helix H12 of chain B forms crystal contacts with chain A's charge clamp of a neighbouring molecule and protrudes from the molecule, as it was observed in other PPARγ structures before (Cronet *et al.*, 2001; Nolte *et al.*, 1998). Helix H12 of chain B thereby occupies the charge clamp into which co-activator peptide binds, thus mimicking the two-turn, amphipathic co-activator α-helix (Figure 2.5-12) (Pochetti *et al.*, 2007). Hence, chain B shows an artificial conformation. Chain A corresponds to the complex structures with co-activator peptides and therefore acts commonly as a model for PPARγ LBD activation. But also the omega loop linking H2' and H3, the most thermally mobile loop of PPARγ's LBD, forms crystal contacts with a related molecule. This loop is stabilised close to the ligand entry side by a hydrogen bond between the carbonyl of Lys265 and the nitrogen of Leu311 of the crystallographic related molecule (Figure 2.5-13A). Thereby the loop occludes parts of the ligand entry site of the inactive form, with His266 and Phe264 pointing into the pocket and to the amorfrutins, leading to an energetically unfavourable position of the ligands. Hence, the amorfrutins delocalise and rotate by almost 180° in the cavity, compared to the active state (Figure 2.5-13B), except amorfrutin 1 that exhibits an alternative conformation in the inactive state.

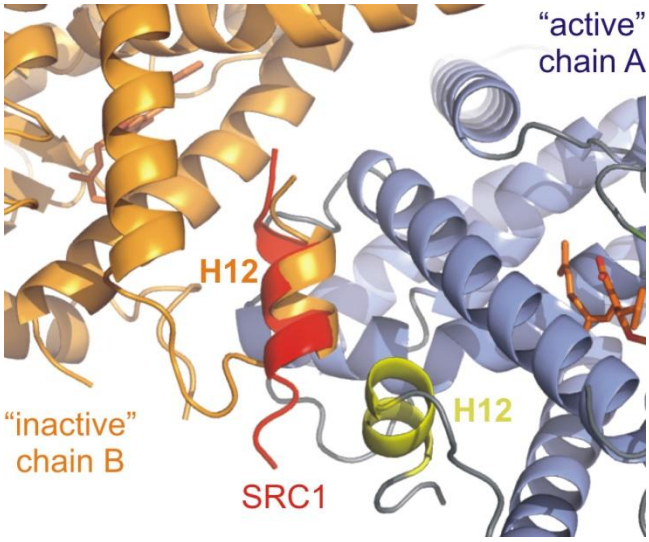

Figure 2.5-12: Crystallographic contacts mimic a co-activator peptide. Cartoon representation of PPARγ's active chain A (blue), interacting with chain B of a symmetry related molecule (orange). Helix H12 of chain B occupies the charge clamp of chain A, thus mimicking a co-activator peptide like SRC1 (red). SRC1 peptide is shown as it lies in the active conformation of the ternary complex with rosiglitazone (PDB 2PRG), generated by superimposition. Helix H12 of active chain A is drawn in yellow. Ribbon diagram is only shown for the complex structure with amorfrutin B.

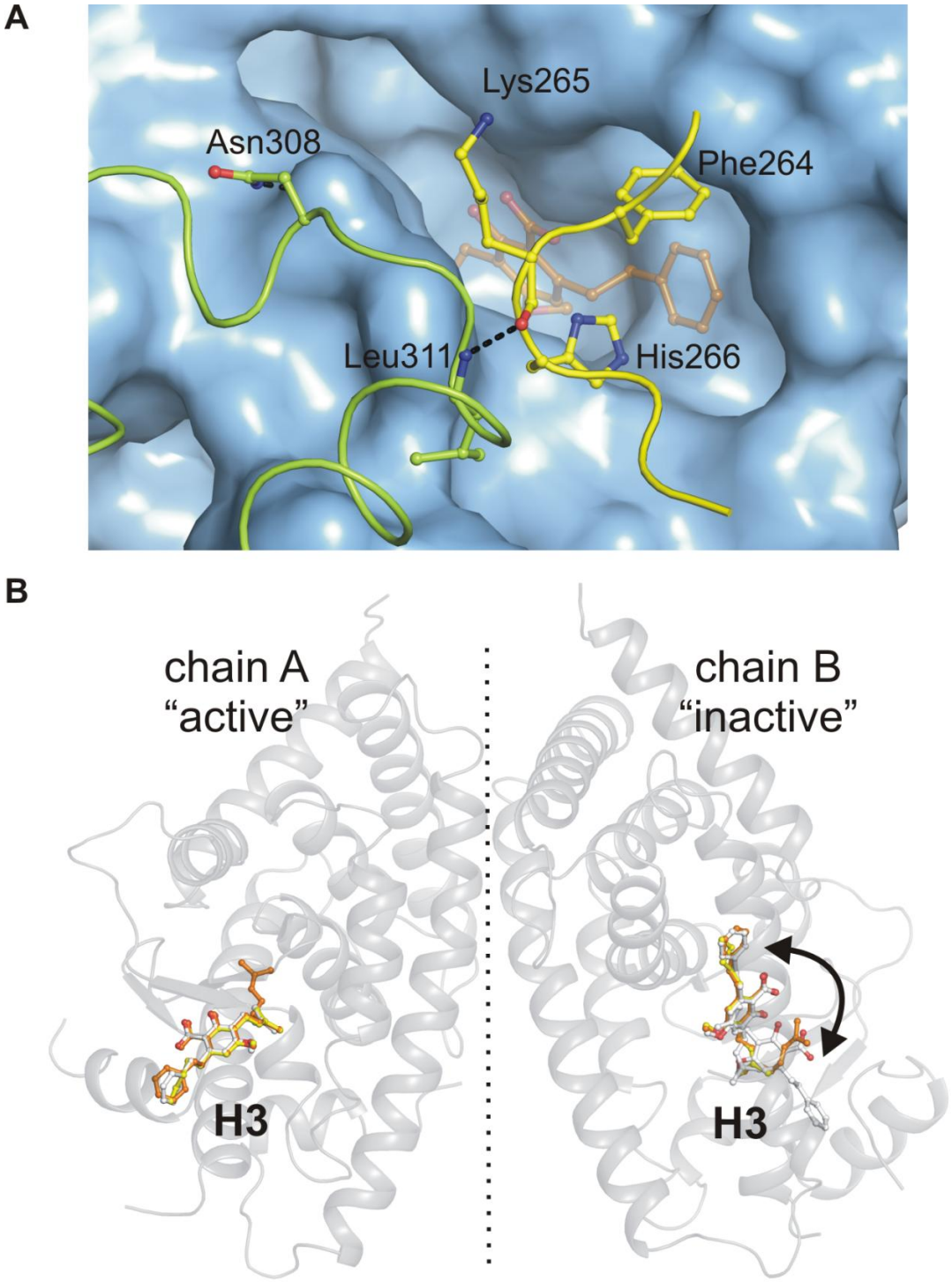

Figure 2.5-13: Crystallographic contacts in the inactive conformation lead to an artificial orientation and localisation of the amorfrutins. (A) surface and ribbon representation of PPARγ`s ligand entry site of the inactive conformation with amorfrutin B (orange) as it lies in the active conformation, generated by superimposition. Interactions of a crystallographically related LBD molecule in a different homodimer (green) stabilise the loop linking H2’ and H3 (yellow) in front of the ligand entry side of the inactive state with His266 and Phe264 pointing into the pocket and to the amorfrutins as they lie in the active conformation. Residues involved are labelled. Oxygen atoms are coloured red and nitrogen atoms blue. Hydrogen bonds are shown as black dashed lines. (B) Localisation of amorfrutin 1 (white), 2 (yellow), and B (orange) in the active and inactive conformation of the LBD homodimer. Note the rotation of the amorfrutins in the inactive conformation by almost 180°, except for amorfrutin 1 (white), showing an alternative conformation. Ribbon diagram is only shown for the amorfrutin B bound form.

2.6 Discussion

2.6.1 Structural insights into amorfrutin recognition by PPARγ

The crystal structures presented here structurally characterise amorfrutin 1, 2, and B, members of a novel class of natural products derived from non-toxic, edible plant roots that possess the ability of a selective PPARγ modulator to separate insulin-sensitisation from adipogenesis (Weidner *et al.,* submitted). In cooperation with the group of Dr. Sascha Sauer (Max Planck Institute for Molecular Genetics, Otto Warburg Laboratory, Berlin, Germany) we identify the three amorfrutins as partial agonists as they show a transcriptional activation of PPARγ below 50% compared to the full agonist rosiglitazone (Weidner *et al.*, submitted) (Table 2.5-2) and interact with the nuclear receptor via a helix H12 independant mechanism by stabilising helix H3 and the β-sheet region, similar to other partial or intermediate agonists of PPARγ (Bruning *et al.*, 2007).

This study shows that the amorfrutin 1, 2, and B bind, due to their structural similarity, with almost identical localisation and orientation close to the ligand entry site of the LBD. All three amorfrutins form extensive van der Waals contacts with the LBD via their isoprenoid, phenyl and aliphatic moieties, most prominently to Ile341 of the β-sheet S3 and Cys285 of helix H3.

All three ligands engage a hydrogen bond network including Arg288 of helix H3 and Ser342 of the β-sheet region. Amorfrutin 1 utilises only its carboxyl group for a direct hydrogen bond to Ser342 and to an alternative conformation of Arg288. Amorfrutin 2 and B utilise also their ortho-hydroxyl group to contact Arg288 and Ser342. However, amorfrutin 2 and B do not directly interact with Arg288 and thus do not induce an alternative conformation of this residue's side chain, but utilize a water-mediated hydrogen bond to Arg288 instead (Figure 2.5-9).

The carboxyl group of the amorfrutins, interacting with the main chain nitrogen of Ser342 and, as shown for amorfrutin 1, the guanidinium group of Arg288, seems to be reasonable for their high affinity. This is supported by a dramatically increased binding constant from 590 nM to 23 µM upon esterification of the acidic head group of the structurally related amorfrutin 5 (Figure 2.6-1) (Weidner *et al.*, submitted).

Regarding the selectivity of amorfrutin 1, 2, and B for PPARγ (60- to 140-fold) over PPARα or PPARδ, Arg288, and in particular its guanidinium group that interacts with the ligands carboxyl or hydroxyl, seems to play the most important role. Arg288 is not conserved among the PPARs. PPARα and PPARδ have a threonine at this position, which cannot interact with the amorfrutins in a similar way to arginine. This suggestion is further supported by a reduced affinity of a naturally occurring PPARγ mutant (R288H) to naturally ligands like the 15-Deoxy-Delta-12,14-prostaglandin J2 (15d-PGJ2) or the fatty acid (9S)-9-Hydroxyoctadeca-10,12-dienoic acid (9(S)-HODE) in comparison to rosiglitazone (Sarraf *et al.*, 1999). Both, 15d-PGJ2 and 9(S)-HODE are shown to use Arg288 for excessive

hydrophobic interactions or hydrogen bondings, respectively (Itoh *et al.*, 2008; Waku *et al.*, 2009), in contrast to the synthetic agonist rosiglitazone (Nolte *et al.*, 1998). Furthermore, hydrogen bonds of the three amorfrutins to the β-sheet region, like the interaction to Ser342, are only dependant on main-chain atoms. Thus these interactions will probably be accomplished by other amino acids, like by the alanine of PPARα and PPARδ at this position. And even Glu343 which is an asparagine or tyrosine in PPARα and PPARδ, respectively, does not seem critical for the recognition, as the side-chain protrudes away from the ligand.

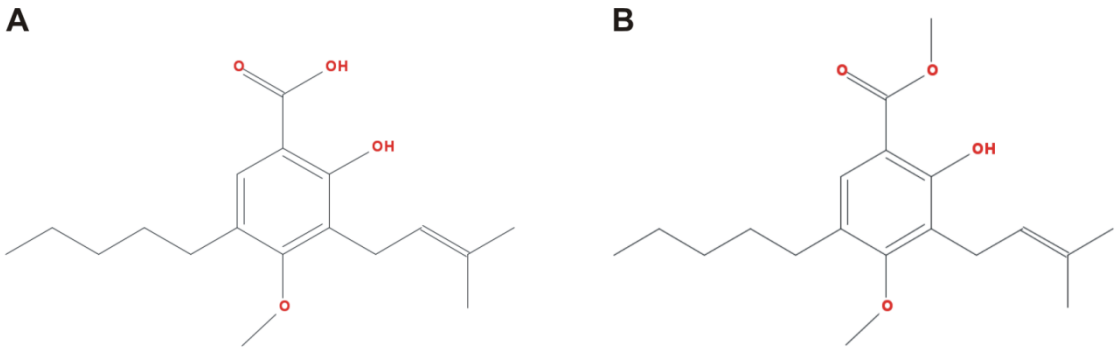

Figure 2.6-1: Esterification of amorfrutin 5. Amorfrutin 5 (A) has a 40-fold higher affinity constant to PPARγ compared to its methylester (B) (Weidner *et al.*, submitted).

2.6.1.1 Structural comparison with other related PPARγ ligands

Comparison of amorfrutin 1, 2, and B with the previously published partial PPARγ agonists nTZDpa and MRL-24 as well as with the intermediate agonist BVT.13 reveals a closely related interaction mode. All these ligands bind at almost the same site and similarly stabilise helix H3 and the β-sheet. Furthermore, all these ligands contain a carboxyl group that directly interacts with the Ser342 of the β-sheet and, with the exception of amorfrutin 1, implicate water molecules for further stabilisation.

The nTZDpa interacts, similar to amorfrutin 1, via its carboxyl group with both, Ser342 and Arg288 in the alternative conformation (Figure 2.5-9A and Figure 2.5-10). Both ligands were described to be capable of avoiding adverse effects like adipogenesis while improving insulin sensitivity in vivo (Berger *et al.*, 2003)(Weidner *et al.*, submitted) and thus have a high potential as drugs to treat type II diabetes without inducing fat storage.

2.6.1.2 Possible mechanism of differential co-regulator recruitment

The structure of the ternary complex of PPARγ, RXRα and DNA revealed three distinct interfaces (Chandra *et al.*, 2008). The DNA binding and ligand-binding domains of the two proteins interact with each other. In addition, the LBD of PPARγ faces with the β-sheet region the DNA binding domain of RXRα (Figure 2.6-2). This β-sheet region is important for DNA binding, as shown by mutating Phe347 of the β-sheet S4 of PPARγ's LBD into alanine, resulting in weakened ability to bind DNA (Chandra *et al.*, 2008). However, rosiglitazone and BVT.13, latter interacting similar to amorfrutins away from helix H12 but involve helix H3 and the LBD's β-sheet region, induce only minor variations of the overall conformation of the ternary complex (Chandra *et al.*, 2008). Thus the effects of amorfrutin 1, 2, and B, as well as the effects of other related ligands on gene transcription are unlikely to be caused by structural changes affecting the DBD of RXRα and its DNA affinity.

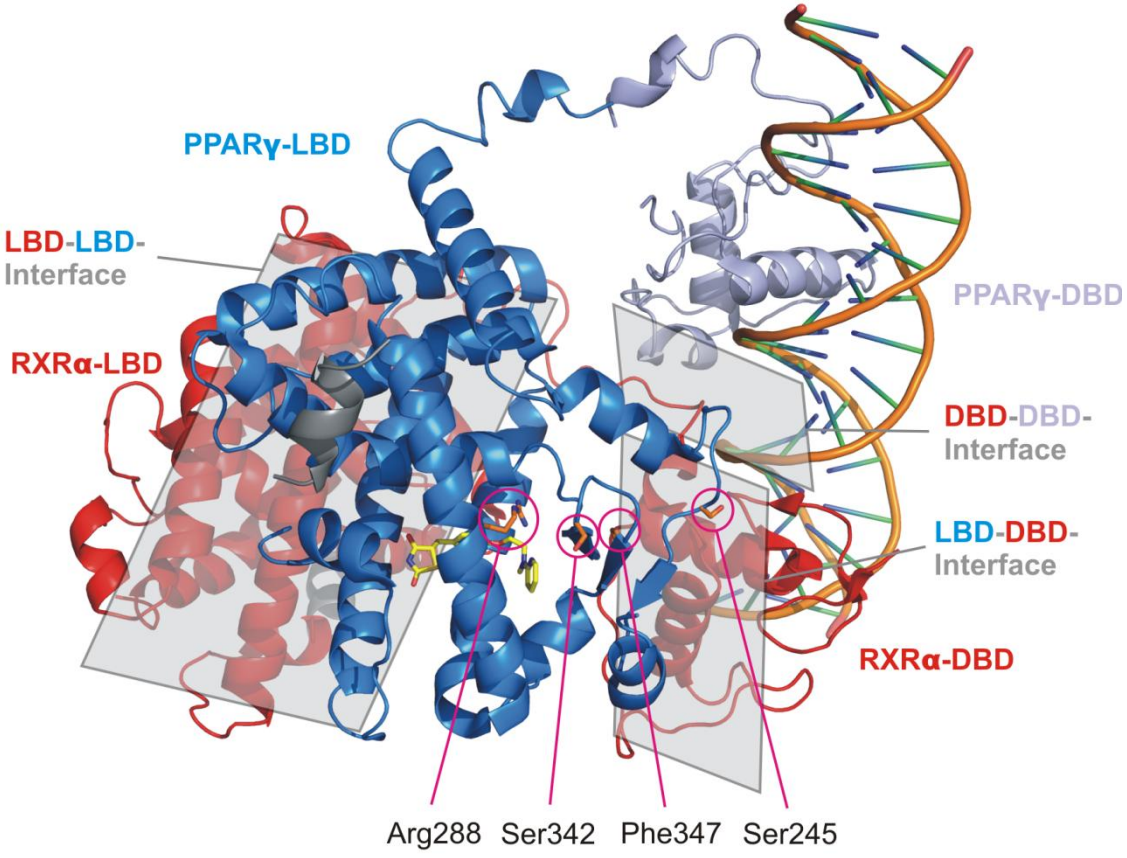

Figure 2.6-2: Overall structure of the PPARγ-RXRα complex on DNA (PDB 3DZY) according to Chandra *et al.* (2008). The expressed LBD of PPARγ for the experimental work in this study is drawn in blue, the PPARγ DBD in light blue, and RXRα in red. Co-activator peptides and rosiglitazone are drawn in black and yellow sticks, respectively. The three different interfaces are indicated by grey planes. The most important residues for amorfrutin interactions, Arg288 and Ser342, as well as the phosphorylation site at Ser245 and the residue Phe347, important for DNA interaction of the heterodimer, are labelled.

PPARγ contains a phosphorylation site at Ser245 (Ser273 in PPARγ2), residing between helix H2 and the β-sheet S1 (Choi *et al.*, 2010), close to RXRα's DBD. High fat feeding induces obesity in mice and activates the protein kinase Cdk5 (cyclin-dependant kinase 5) in adipose tissue, resulting in phosphorylation of this residue (Choi *et al.*, 2010). The phosphorylation has no influence on DNA binding or general transcriptional activity, but alters the expression of a subset of genes with regulatory functions in metabolism, including a reduced transcription of the insulin-sensitising adipokine, adiponectin (Choi *et al.*, 2010). Amorfrutin 1 and rosiglitazone were shown to block the phosphorylation of Ser245, which is apparently closely associated with their anti-diabetic effects (Choi *et al.*, 2010) (Weidner *et al.*, submitted). It is hypothesized that ligand-mediated interactions between helix H3 and the β-sheet 'freezes' this region in a conformation less favourable to Cdk5 phosphorylation (Choi *et al.*, 2010). This leads to the suggestion of differentially regulated co-regulator recruitment in a phosphorylation-dependant manner (Choi *et al.*, 2010) by a novel co-activator binding surface close to Ser245. Thus the inhibited phosphorylation by rosiglitazone or amorfrutin 1 may initiate the recruitment of additional co-regulators, reasonable for dysregulated gene activation, like the enhanced transcription of the insulin-sensitising adiponectin, and the increased insulin sensitivity. Furthermore, regions outside the co-activators LXXLL motifs may also be involved and affect co-activator recognition and regulation (Klein *et al.*, 2005).

2.6.1.3 The active and inactive state of the PPARγ's LBD

The data demonstrate the unreliability of the inactive conformation even for the helix H12 independant PPARγ activation. Crystallographic contacts to the loop linking H2' and H3, called the omega-loop, lead to rotation and relocation of the amorfrutins in the cavity (section 2.5.4). The omega-loop seems to be responsible for modulation of PPARγ activity by endogenous fatty acids like 15d-PGJ2 (Waku *et al.*, 2009), illustrating the importance to distinguish between active and inactive form. Furthermore, crystallographic contacts to the omega-loop occlude parts of the ligand entry site and may thus exacerbate soakings of different ligands.

2.7 Outlook

Due to the reported insulin-sensitising effects, the reduced fat storage and the improvement of metabolic and inflammatory parameters, which were ascribed to amorfrutin 1, it would be an exciting task to study the different effects of amorfrutin B, which exhibits a 12-fold higher binding affinity to PPARγ and to test if this compound exhibits improved, similar, or probably reduced effects. Especially the effect on the Cdk5 mediated phosphorylation of PPARγ's LBD at Ser245 would be of special interest, as, to the current knowledge, it seems to be strongly associated with the anti-diabetic effect of substances like rosiglitazone, MRL-24, or amorfrutin 1. The higher binding affinity of amorfrutin B may contribute to a better stabilisation of the receptor and probably rigidify it in a higher degree as amorfrutin 1 or rosiglitazone do and hence may protect Ser245 more efficiently from Cdk5 mediated phosphorylation. It would be also possible that a structural related ligand with a higher binding affinity may reduce the dosing for the desired effect, which is currently with 10 mg/kg/d amorfrutin 1 (Weidner *et al.*, submitted) quite high. On the contrary, the more lipophilic structure and a possibly resulting reduced bioavailability in the organism may lead to reduced effects and a reduced insulin sensitising action. These effects remain elusive and have to be discovered in the future.

Furthermore, as the PPARγ ligand 15d-PGJ2, a naturally occurring prostaglandin, was described to modify the receptors transcriptional activity by inducing structural changes to the omega-loop (Waku *et al.*, 2009), residing between helix H2' and H3, it would be also an exciting task to test if amorfrutins also induce a comparable and transcriptional relevant change to this region. As this region exhibits a high flexibility and thus was not visible in the present crystal structures, biophysical techniques appear to be more suitable to address this question. Method of choice may be the measurement of tryptophan fluorescence spectra of a K265W mutant of PPARγ (Waku *et al.*, 2009). The mutation seems to be uncritical for ligand binding as K265 protrudes away from the LBD. The residue resides in the omega-loop and hence environmental changes around this area should become visible by changes of the spectra. Furthermore, the mutants I267A and F287A, residues of the omega-loop and helix H3, respectively, were described to dramatically decrease the receptors transcriptional activity in response to 15d-PGJ2, as these residues are responsible for the induced structural changes of the omega-loop without disturbing 15d-PGJ2 affinity (Waku *et al.*, 2009). The amorfrutins interact close to the ligand entry side and the omega-loop and thus probably also affect the omega-loop conformation. Thus it would be of interest to analyse the amorfrutin mediated cell based transcriptional activity of these mutants in comparison to the wilde-type.

Additional mutants like R288T, R288A and R288H may be also of interest to quantify the contribution of the LBD's Arg288 to amorfrutin binding, transcriptional activity, and to the selectivity of PPARγ to the amorfrutins in comparison to the other PPAR subtypes. R288T should mimic the residue of PPARα and PPARδ at this position and thus may give insights in PPARγ's amorfrutin selectivity, compared with the wilde-type and the R288A mutant. The naturally occurring PPARγ mutant R288H is suggested to be associated with human colorectal cancer and was described to exhibit reduced affinity to naturally occurring ligands like 15d-PGJ2 or 9(S)-HODE (Section 2.6.1), shown to excessively interact with Arg288 (Sarraf *et al.*, 1999) (Section 2.6.1).

However, one of the most important questions regarding PPARγ mediated transcriptional activation or dysregulation is addressed to probably additional co-regulator recruitment. The omega-loop region and the region around the phosphorylation site at Ser245 appear currently as the most suitable regions for such co-regulator recruitment. Thus it would be an exciting task to identify possible new co-regulators of PPARγ and discover new aspects of the mechanism of nuclear receptor regulation.

Anyway, the insulin sensitising effect and the reduced fat storage upon ingestion of the non-toxic amorfrutin 1 makes this naturally compound already very attractive for the possible usage as a drug to treat diseases like type II diabetes and/or obesity. Furthermore, as the amorfrutins are natural and non-toxic compounds, the usage of amorfrutins as dietary supplements to prevent type II diabetes and/or obesity may be also possible. Thus these compounds need to be tested by a human study to unveil possible human specific adverse effects and to give more detailed information about the metabolic action of these substances.

3 References

Adams, M., Reginato, M.J., Shao, D., Lazar, M.A., and Chatterjee, V.K. (1997). Transcriptional activation by peroxisome proliferator-activated receptor gamma is inhibited by phosphorylation at a consensus mitogen-activated protein kinase site. *J Biol Chem* **272**, 5128-5132.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.

Ahmed, S., Goh, W.I., and Bu, W. (2009). I-BAR domains, IRSp53 and filopodium formation. *Semin Cell Dev Biol* **21**, 350-356.

Arold, S., O'Brien, R., Franken, P., Strub, M.P., Hoh, F., Dumas, C., and Ladbury, J.E. (1998). RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. *Biochemistry* **37**, 14683-14691.

Bach, R.R. (1988). Initiation of coagulation by tissue factor. *CRC Crit Rev Biochem* **23**, 339-368.

Ball, L.J., Kühne, R., Schneider-Mergener, J., and Oschkinat, H. (2005). Recognition of Proline-Rich Motifs by Protein–Protein-Interaction Domains. *Angew Chem Int Ed* **44**, 2852-2869.

Bast, D.J., Banerjee, L., Clark, C., Read, R.J., and Brunton, J.L. (1999). The identification of three biologically relevant globotriaosyl ceramide receptor binding sites on the Verotoxin 1 B subunit. *Mol Microbiol* **32**, 953-960.

Berger, J., and Moller, D.E. (2002). The mechanisms of action of PPARs. *Annu Rev Med* **53**, 409-435.

Berger, J.P., Petro, A.E., Macnaul, K.L., Kelly, L.J., Zhang, B.B., Richards, K., Elbrecht, A., Johnson, B.A., Zhou, G., Doebber, T.W., *et al.* (2003). Distinct properties and advantages of a novel peroxisome proliferator-activated protein [gamma] selective modulator. *Mol Endocrinol* **17**, 662-676.

Bergfors, T.M. (1999). Protein Crystallization: Techniques, Strategies, and Tips. In, T.M. Bergfors, ed. (La Jolla, International University Line), pp. 29-38.

Bockemühl, J., Karch, H., and Tschäpe, H. (1997). Infektionen des Menschen durch enterohämorrhagische *Escherichia coli* (EHEC) in Deutschland, 1996. *Bundesgesundheitsblatt*, 194-197.

Bockmann, J., Kreutz, M.R., Gundelfinger, E.D., and Bockers, T.M. (2002). ProSAP/Shank postsynaptic density proteins interact with insulin receptor tyrosine kinase substrate IRSp53. *J Neurochem* **83**, 1013-1017.

Brady, M.J., Campellone, K.G., Ghildiyal, M., and Leong, J.M. (2007). Enterohaemorrhagic and enteropathogenic Escherichia coli Tir proteins trigger a common Nck-independent actin assembly pathway. *Cell Microbiol* **9**, 2242-2253.

Bruning, J.B., Chalmers, M.J., Prasad, S., Busby, S.A., Kamenecka, T.M., He, Y., Nettles, K.W., and Griffin, P.R. (2007). Partial agonists activate PPARgamma using a helix 12 independent mechanism. *Structure* **15**, 1258-1271.

Campellone, K.G., Brady, M.J., Alamares, J.G., Rowe, D.C., Skehan, B.M., Tipper, D.J., and Leong, J.M. (2006). Enterohaemorrhagic Escherichia coli Tir requires a C-terminal 12-residue peptide to initiate $EspF_U$-mediated actin assembly and harbours N-terminal sequences that influence pedestal length. *Cell Microbiol* **8**, 1488-1503.

Campellone, K.G., and Leong, J.M. (2003). Tails of two Tirs: actin pedestal formation by enteropathogenic E. coli and enterohemorrhagic E. coli O157:H7. *Curr Opin Microbiol* **6**, 82-90.

Campellone, K.G., Robbins, D., and Leong, J.M. (2004). EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell* **7**, 217-228.

Castillo, G., Brun, R.P., Rosenfield, J.K., Hauser, S., Park, C.W., Troy, A.E., Wright, M.E., and Spiegelman, B.M. (1999). An adipogenic cofactor bound by the differentiation domain of PPARgamma. *EMBO J* **18**, 3676-3687.

Chandra, V., Huang, P., Hamuro, Y., Raghuram, S., Wang, Y., Burris, T.P., and Rastinejad, F. (2008). Structure of the intact PPAR-gamma-RXR-alpha nuclear receptor complex on DNA. *Nature*, 350-356.

Chen, J.D., and Evans, R.M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377**, 454-457.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21.

Cheng, H.C., Skehan, B.M., Campellone, K.G., Leong, J.M., and Rosen, M.K. (2008). Structural mechanism of WASP activation by the enterohaemorrhagic E. coli effector EspF(U). *Nature* **454**, 1009-1013.

Choi, J.H., Banks, A.S., Estall, J.L., Kajimura, S., Bostrom, P., Laznik, D., Ruas, J.L., Chalmers, M.J., Kamenecka, T.M., Bluher, M., *et al.* (2010). Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPARgamma by Cdk5. *Nature* **466**, 451-456.

Cronet, P., Petersen, J.F., Folmer, R., Blomberg, N., Sjoblom, K., Karlsson, U., Lindstedt, E.L., and Bamberg, K. (2001). Structure of the PPARalpha and -gamma ligand binding domain in complex with AZ 242; ligand selectivity and agonist activation in the PPAR family. *Structure* **9**, 699-706.

Croxen, M.A., and Finlay, B.B. (2010). Molecular mechanisms of Escherichia coli pathogenicity. *Nat Rev Microbiol* **8**, 26-38.

Darimont, B.D., Wagner, R.L., Apriletti, J.W., Stallcup, M.R., Kushner, P.J., Baxter, J.D., Fletterick, R.J., and Yamamoto, K.R. (1998). Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev* **12**, 3343-3356.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S.*, et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids*. Nucleic Acids Res* **35**, W375-383.

de Beer, T., Hoofnagle, A.N., Enmon, J.L., Bowers, R.C., Yamabhai, M., Kay, B.K., and Overduin, M. (2000). Molecular mechanism of NPF recognition by EH domains*. Nat Struct Mol Biol* **7**, 1018-1022.

de Groot, J.C., Schlüter, K., Carius, Y., Quedenau, C., Vingadassalom, D., Faix, J., Weiss, S.M., Reichelt, J., Standfuss-Gabisch, C., Lesser, C.F.*, et al.* (2011). Structural Basis for Complex Formation between Human IRSp53 and the Translocated Intimin Receptor Tir of Enterohemorrhagic E. coli*. Structure* **19**, 1294-1306.

Dean, P., Maresca, M., Schuller, S., Phillips, A.D., and Kenny, B. (2006). Potent diarrheagenic mechanism mediated by the cooperative action of three enteropathogenic Escherichia coli-injected effector proteins*. PNAS* **103**, 1876-1881.

Disanza, A., Mantoani, S., Hertzog, M., Gerboth, S., Frittoli, E., Steffen, A., Berhoerster, K., Kreienkamp, H.-J., Milanesi, F., Fiore, P.P.D.*, et al.* (2006). Regulation of cell shape by Cdc42 is mediated by the synergic actin-bundling activity of the Eps8-IRSp53 complex*. Nat Cell Biol* **8**, 1337-1347.

Duguid, J.P., Smith, I.W., Dempster, G., and Edmunds, P.N. (1955). Non-flagellar filamentous appendages (fimbriae) and haemagglutinating activity in Bacterium coli*. J Pathol Bacteriol* **70**, 335-348.

Eck, M.J., Dhe-Paganon, S., Trüb, T., Nolte, R.T., and Shoelson, S.E. (1996). Structure of the IRS-1 PTB Domain Bound to the Juxtamembrane Region of the Insulin Receptor*. Cell* **85**, 695-705.

Enami, M., Nakasone, N., Honma, Y., Kakinohana, S., Kudaka, J., and Iwanaga, M. (1999). Expression of type I pili is abolished in verotoxin-producing Escherichia coli O157*. FEMS Microbiol Lett* **179**, 467-472.

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T., and Igarashi, K. (1988). Site of action of a Vero toxin (VT2) from Escherichia coli O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins*. Eur J Biochem* **171**, 45-50.

Esmaili, A., Nazir, S.F., Borthakur, A., Yu, D., Turner, J.R., Saksena, S., Singla, A., Hecht, G.A., Alrefai, W.A., and Gill, R.K. (2009). Enteropathogenic Escherichia coli infection inhibits intestinal serotonin transporter function and expression*. Gastroenterology* **137**, 2074-2083.

Evans, R.M. (1988). The steroid and thyroid hormone receptor superfamily*. Science* **240**, 889-895.

Frankel, G., and Phillips, A.D. (2008). Attaching effacing Escherichia coli and paradigms of Tir-triggered actin polymerization: getting off the pedestal*. Cell Microbiol* **10**, 549-556.

Fraser, M.E., Chernaia, M.M., Kozlov, Y.V., and James, M.N. (1994). Crystal structure of the holotoxin from Shigella dysenteriae at 2.5 A resolution*. Nat Struct Biol* **1**, 59-64.

Frese, S., Schubert, W.-D., Findeis, A.C., Marquardt, T., Roske, Y.S., Stradal, T.E.B., and Heinz, D.W. (2006). The Phosphotyrosine Peptide Binding Specificity of Nck1 and Nck2 Src Homology 2 Domains*. J Biol Chem* **281**, 18236-18245.

Frost, A., Unger, V.M., and De Camilli, P. (2009). The BAR domain superfamily: membrane-molding macromolecules. *Cell* **137**, 191-196.

Fuchs, G., Mobassaleh, M., Donohue-Rolfe, A., Montgomery, R.K., Grand, R.J., and Keusch, G.T. (1986). Pathogenesis of Shigella diarrhea: rabbit intestinal cell microvillus membrane binding site for Shigella toxin. *Infect Immun* **53**, 372-377.

Fujiwara, T., Mammoto, A., Kim, Y., and Takai, Y. (2000). Rho small G-protein-dependent binding of mDia to an Src homology 3 domain-containing IRSp53/BAIAP2. *Biochem Biophys Res Commun* **271**, 626-629.

Garmendia, J., Frankel, G., and Crepin, V.F. (2005). Enteropathogenic and enterohemorrhagic Escherichia coli infections: translocation, translocation, translocation. *Infect Immun* **73**, 2573-2585.

Garmendia, J., Phillips, A.D., Carlier, M.F., Chong, Y., Schuller, S., Marches, O., Dahan, S., Oswald, E., Shaw, R.K., Knutton, S., *et al.* (2004). TccP is an enterohaemorrhagic Escherichia coli O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cell Microbiol* **6**, 1167-1183.

Gill, R.K., Borthakur, A., Hodges, K., Turner, J.R., Clayburgh, D.R., Saksena, S., Zaheer, A., Ramaswamy, K., Hecht, G., and Dudeja, P.K. (2007). Mechanism underlying inhibition of intestinal apical Cl/OH exchange following infection with enteropathogenic E. coli. *J Clin Invest* **117**, 428-437.

Glass, C.K., and Ogawa, S. (2006). Combinatorial roles of nuclear receptors in inflammation and immunity. *Nat Rev Immunol* **6**, 44-55.

Gobert, A.P., Vareille, M., Glasser, A.L., Hindre, T., de Sablet, T., and Martin, C. (2007). Shiga toxin produced by enterohemorrhagic Escherichia coli inhibits PI3K/NF-kappaB signaling pathway in globotriaosylceramide-3-negative human intestinal epithelial cells. *J Immunol* **178**, 8168-8174.

Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucl Acids Res* **31**, 3320-3323.

Gruenheid, S., DeVinney, R., Bladt, F., Goosney, D., Gelkop, S., Gish, G.D., Pawson, T., and Finlay, B.B. (2001). Enteropathogenic E. coli Tir binds Nck to initiate actin pedestal formation in host cells. *Nat Cell Biol* **3**, 856-859.

Gruppo, R.A., and Rother, R.P. (2009). Eculizumab for congenital atypical hemolytic-uremic syndrome. *N Engl J Med* **360**, 544-546.

Guttman, J.A., Lin, A.E., Li, Y., Bechberger, J., Naus, C.C., Vogl, A.W., and Finlay, B.B. (2011). Gap junction hemichannels contribute to the generation of diarrhoea during infectious enteric disease. *Gut* **59**, 218-226.

Guttman, J.A., Samji, F.N., Li, Y., Deng, W., Lin, A., and Finlay, B.B. (2007). Aquaporins contribute to diarrhoea caused by attaching and effacing bacterial pathogens. *Cell Microbiol* **9**, 131-141.

Harrison, L.M., van Haaften, W.C., and Tesh, V.L. (2004). Regulation of proinflammatory cytokine expression by Shiga toxin 1 and/or lipopolysaccharides in the human monocytic cell line THP-1. *Infect Immun* **72**, 2618-2627.

Hayward, R.D., Leong, J.M., Koronakis, V., and Campellone, K.G. (2006). Exploiting pathogenic Escherichia coli to model transmembrane receptor signalling. *Nat Rev Microbiol* **4**, 358-370.

Heery, D.M., Kalkhoven, E., Hoare, S., and Parker, M.G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733-736.

Hertzke, D.M., Cowan, L.A., Schoning, P., and Fenwick, B.W. (1995). Glomerular ultrastructural lesions of idiopathic cutaneous and renal glomerular vasculopathy of greyhounds. *Vet Pathol* **32**, 451-459.

Hodges, K., Alto, N.M., Ramaswamy, K., Dudeja, P.K., and Hecht, G. (2008). The enteropathogenic Escherichia coli effector protein EspF decreases sodium hydrogen exchanger 3 activity. *Cell Microbiol* **10**, 1735-1745.

Horlein, A.J., Naar, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C.K., *et al.* (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377**, 397-404.

Hu, E., Kim, J.B., Sarraf, P., and Spiegelman, B.M. (1996). Inhibition of adipogenesis through MAP kinase-mediated phosphorylation of PPARgamma. *Science* **274**, 2100-2103.

Hu, X., and Lazar, M.A. (1999). The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. *Nature* **402**, 93-96.

Huang, Z., Sutton, S.E., Wallenfang, A.J., Orchard, R.C., Wu, X., Feng, Y., Chai, J., and Alto, N.M. (2009). Structural insights into host GTPase isoform selection by a family of bacterial GEF mimics. *Nat Struct Mol Biol* **16**, 853-860.

Institut Pasteur (2007). In Antigen Formulae of the *Salmonella* Serovars (Paris, WHO Collaborating Center for Reference and Research on Salmonella).

Itoh, T., Fairall, L., Amin, K., Inaba, Y., Szanto, A., Balint, B.L., Nagy, L., Yamamoto, K., and Schwabe, J.W. (2008). Structural basis for the activation of PPARgamma by oxidized fatty acids. *Nat Struct Mol Biol* **15**, 924-931.

Jackson, M.P., Newland, J.W., Holmes, R.K., and O'Brien, A.D. (1987). Nucleotide sequence analysis of the structural genes for Shiga-like toxin I encoded by bacteriophage 933J from Escherichia coli. *Microb Pathog* **2**, 147-153.

Johannes, L., and Romer, W. (2010). Shiga toxins-from cell biology to biomedical applications. *Nat Rev Microbiol* **8**, 105-116.

Johnson, J.R. (1991). Virulence factors in Escherichia coli urinary tract infection. *Clin Microbiol Rev* **4**, 80-128.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* **26**, 795-800.

Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004). Pathogenic Escherichia coli. *Nat Rev Microbiol* **2**, 123-140.

Karch, H. (2001). The role of virulence factors in enterohemorrhagic Escherichia coli (EHEC)--associated hemolytic-uremic syndrome. *Semin Thromb Hemost* **27**, 207-213.

Kardinal, C., Konkol, B., Schulz, A., Posern, G., Lin, H., Adermann, K., Eulitz, M., Estrov, Z., Talpaz, M., Arlinghaus, R.B.*, et al.* (2000). Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients*. Faseb J* **14**, 1529-1538.

Karmali, M.A., Petric, M., Lim, C., Fleming, P.C., Arbus, G.S., and Lior, H. (1985). The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing Escherichia coli*. J Infect Dis* **151**, 775-782.

Kenny, B., Ellis, S., Leard, A.D., Warawa, J., Mellor, H., and Jepson, M.A. (2002). Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic Escherichia coli effector molecules*. Mol Microbiol* **44**, 1095-1107.

Kirsch, P., Jores, J., and Wieler, L.H. (2004). [Plasticity of bacterial genomes: pathogenicity islands and the locus of enterocyte effacement (LEE)]*. Berl Munch Tierarztl Wochenschr* **117**, 116-129.

Klein, F.A., Atkinson, R.A., Potier, N., Moras, D., and Cavarelli, J. (2005). Biochemical and NMR mapping of the interface between CREB-binding protein and ligand binding domains of nuclear receptor: beyond the LXXLL motif*. J Biol Chem* **280**, 5682-5692.

Kliewer, S.A., Sundseth, S.S., Jones, S.A., Brown, P.J., Wisely, G.B., Koble, C.S., Devchand, P., Wahli, W., Willson, T.M., Lenhard, J.M.*, et al.* (1997). Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors alpha and gamma*. PNAS* **94**, 4318-4323.

Kodera, Y., Takeyama, K., Murayama, A., Suzawa, M., Masuhiro, Y., and Kato, S. (2000). Ligand type-specific interactions of peroxisome proliferator-activated receptor gamma with transcriptional coactivators*. J Biol Chem* **275**, 33201-33204.

Krugmann, S., Jordens, I., Gevaert, K., Driessens, M., Vandekerckhove, J., and Hall, A. (2001). Cdc42 induces filopodia by promoting the formation of an IRSp53:Mena complex*. Curr Biol* **11**, 1645-1655.

Kuriyan, J., and Cowburn, D. (1997). Modular peptide recognition domains in eukaryotic signaling*. Annu Rev Biophys Biomol Struct* **26**, 259-288.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature* **227**, 680-685.

Lapeyraque, A.L., Malina, M., Fremeaux-Bacchi, V., Boppel, T., Kirschfink, M., Oualha, M., Proulx, F., Clermont, M.J., Le Deist, F., Niaudet, P.*, et al.* (2011). Complement Blockade in Severe Shiga-Toxin-Associated HUS*. N Engl J Med*.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0*. Bioinformatics* **23**, 2947-2948.

Leavitt, S., and Freire, E. (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry*. Curr Opin Struct Biol* **11**, 560-566.

Lee, S.H., Kerff, F., Chereau, D., Ferron, F., Klug, A., and Dominguez, R. (2007). Structural basis for the actin-binding function of missing-in-metastasis*. Structure* **15**, 145-155.

Li, Y., Kovach, A., Suino-Powell, K., Martynowski, D., and Xu, H.E. (2008). Structural and biochemical basis for the binding selectivity of peroxisome proliferator-activated receptor gamma to PGC-1alpha*. J Biol Chem* **283**, 19132-19139.

Lim, J.P., and Gleeson, P.A. (2011). Macropinocytosis: an endocytic pathway for internalising large gulps. *Immunol Cell Biol*.

Lim, K.B., Bu, W., Goh, W.I., Koh, E., Ong, S.H., Pawson, T., Sudhaharan, T., and Ahmed, S. (2008). The Cdc42 Effector IRSp53 Generates Filopodia by Coupling Membrane Protrusion with Actin Dynamics. *J Biol Chem* **283**, 20454-20472.

Lingwood, C.A. (1994). Verotoxin-binding in human renal sections. *Nephron* **66**, 21-28.

Lingwood, C.A., Binnington, B., Manis, A., and Branch, D.R. (2010a). Globotriaosyl ceramide receptor function - where membrane structure and pathology intersect. *FEBS Lett* **584**, 1879-1886.

Lingwood, C.A., Manis, A., Mahfoud, R., Khan, F., Binnington, B., and Mylvaganam, M. (2010b). New aspects of the regulation of glycosphingolipid receptor function. *Chem Phys Lipids* **163**, 27-35.

London, N., Movshovitz-Attias, D., and Schueler-Furman, O. (2010). The Structural Basis of Peptide-Protein Binding Strategies. *Structure* **18**, 188-199.

Luo, Y., Frey, E.A., Pfuetzner, R.A., Creagh, A.L., Knoechel, D.G., Haynes, C.A., Finlay, B.B., and Strynadka, N.C.J. (2000). Crystal structure of enteropathogenic Escherichia coli intimin-receptor complex. *Nature* **405**, 1073-1077.

Ma, C., Wickham, M.E., Guttman, J.A., Deng, W., Walker, J., Madsen, K.L., Jacobson, K., Vogl, W.A., Finlay, B.B., and Vallance, B.A. (2006). Citrobacter rodentium infection causes both mitochondrial dysfunction and intestinal epithelial barrier disruption in vivo: role of mitochondrial associated protein (Map). *Cell Microbiol* **8**, 1669-1686.

Mahfoud, R., Manis, A., Binnington, B., Ackerley, C., and Lingwood, C.A. (2010). A major fraction of glycosphingolipids in model and cellular cholesterol-containing membranes is undetectable by their binding proteins. *J Biol Chem* **285**, 36049-36059.

Malyukova, I., Murray, K.F., Zhu, C., Boedeker, E., Kane, A., Patterson, K., Peterson, J.R., Donowitz, M., and Kovbasnjuk, O. (2009). Macropinocytosis in Shiga toxin 1 uptake by human intestinal epithelial cells and transcellular transcytosis. *Am J Physiol Gastrointest Liver Physiol* **296**, G78-92.

Mattila, P.K., Pykäläinen, A., Saarikangas, J., Paavilainen, V.O., Vihinen, H., Jokitalo, E., and Lappalainen, P. (2007). Missing-in-metastasis and IRSp53 deform $PI(4,5)P_2$-rich membranes by an inverse BAR domain-like mechanism. *J Cell Biol* **176**, 953-964.

Mayer, B.J. (2001). SH3 domains: complexity in moderation. *J Cell Sci* **114**, 1253-1263.

Mayer, B.J., and Saksela, K. (2004). SH3 Domains. In Modular Protein Domains, G. Cesareni, M. Gimona, M. Sudol, and M. Yaffe, eds. (Weinheim, Willey-VCH Verlag GmbH), pp. 37-85.

McLachlan, A.D. (1982). Rapid comparison of protein structures. *Acta Crystallogr A* **38**, 871-873.

Mead, P.S., and Griffin, P.M. (1998). Escherichia coli O157:H7. *Lancet* **352**, 1207-1212.

Meinild, A., Klaerke, D.A., Loo, D.D., Wright, E.M., and Zeuthen, T. (1998). The human Na+-glucose cotransporter is a molecular water pump. *J Physiol* **508**, 15-21.

Michalik, L., Desvergne, B., Tan, N.S., Basu-Modak, S., Escher, P., Rieusset, J., Peters, J.M., Kaya, G., Gonzalez, F.J., Zakany, J., *et al.* (2001). Impaired skin wound healing in peroxisome proliferator-activated receptor (PPAR)alpha and PPARbeta mutant mice. *J Cell Biol* **154**, 799-814.

Miki, H., Yamaguchi, H., Suetsugu, S., and Takenawa, T. (2000). IRSp53 is an essential intermediate between Rac and WAVE in the regulation of membrane ruffling. *Nature* **408**, 732-735.

Millard, T.H., Bompard, G., Heung, M.Y., Dafforn, T.R., Scott, D.J., Machesky, L.M., and Fütterer, K. (2005). Structural basis of filopodia formation induced by the IRSp53/MIM homology domain of human IRSp53. *EMBO J* **24**, 240-250.

Mukherjee, J., Chios, K., Fishwild, D., Hudson, D., O'Donnell, S., Rich, S.M., Donohue-Rolfe, A., and Tzipori, S. (2002). Human Stx2-specific monoclonal antibodies prevent systemic complications of Escherichia coli O157:H7 infection. *Infect Immun* **70**, 612-619.

Mulvey, G.L., Marcato, P., Kitov, P.I., Sadowska, J., Bundle, D.R., and Armstrong, G.D. (2003). Assessment in mice of the therapeutic potential of tailored, multivalent Shiga toxin carbohydrate ligands. *J Infect Dis* **187**, 640-649.

Murphy, G.J., and Holder, J.C. (2000). PPAR-gamma agonists: therapeutic role in diabetes, inflammation and cancer. *Trends Pharmacol Sci* **21**, 469-474.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**, 240-255.

Nagy, L., Kao, H.Y., Love, J.D., Li, C., Banayo, E., Gooch, J.T., Krishna, V., Chatterjee, K., Evans, R.M., and Schwabe, J.W. (1999). Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes Dev* **13**, 3209-3216.

Nagy, L., and Schwabe, J.W. (2004). Mechanism of the nuclear receptor molecular switch. *Trends Biochem Sci* **29**, 317-324.

Nishikawa, K., Matsuoka, K., Kita, E., Okabe, N., Mizuguchi, M., Hino, K., Miyazawa, S., Yamasaki, C., Aoki, J., Takashima, S., *et al.* (2002). A therapeutic agent with oriented carbohydrates for treatment of infections by Shiga toxin-producing Escherichia coli O157:H7. *PNAS* **99**, 7669-7674.

Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Willson, T.M., Glass, C.K., and Milburn, M.V. (1998). Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* **395**, 137-143.

Noris, M., and Remuzzi, G. (2005). Hemolytic uremic syndrome. *J Am Soc Nephrol* **16**, 1035-1050.

Novy, R., Drott, D., Yaeger, K., and Mierendorf, R. (2001). Overcoming the codon bias of E. coli for enhanced protein expression. *inNovations* **12**, 1-3.

Nutikka, A., and Lingwood, C. (2004). Generation of receptor-active, globotriaosyl ceramide/cholesterol lipid 'rafts' in vitro : A new assay to define factors affecting glycosphingolipid receptor activity. *Glycoconj J* **20**, 33-38.

O'Brien, A.D., Newland, J.W., Miller, S.F., Holmes, R.K., Smith, H.W., and Formal, S.B. (1984). Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694-696.

O'Brien, A.D., Tesh, V.L., Donohue-Rolfe, A., Jackson, M.P., Olsnes, S., Sandvig, K., Lindberg, A.A., and Keusch, G.T. (1992). Shiga toxin: biochemistry, genetics, mode of action, and role in pathogenesis. *Curr Top Microbiol Immunol* **180**, 65-94.

Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. (2003). Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucl Acids Res* **31**, 3635-3641.

Obrig, T.G. (2010). Escherichia coli Shiga Toxin Mechanisms of Action in Renal Disease. *Toxins (Basel)* **2**, 2769-2794.

Obrig, T.G., Del Vecchio, P.J., Brown, J.E., Moran, T.P., Rowland, B.M., Judge, T.K., and Rothman, S.W. (1988). Direct cytotoxic action of Shiga toxin on human vascular endothelial cells. *Infect Immun* **56**, 2373-2378.

Okamura-Oho, Y., Miyashita, T., Ohmi, K., and Yamada, M. (1999). Dentatorubral-pallidoluysian atrophy protein interacts through a proline-rich region near polyglutamine with the SH3 domain of an insulin receptor tyrosine kinase substrate. *Hum Mol Genet* **8**, 947-957.

Orskov, F., and Orskov, I. (1992). Escherichia coli serotyping and disease in man and animals. *Can J Microbiol* **38**, 699-704.

Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-450.

Parker, C.J., Kar, S., and Kirkpatrick, P. (2007). Eculizumab. *Nat Rev Drug Discov* **6**, 515-516.

Paton, A.W., Morona, R., and Paton, J.C. (2000). A new biological agent for treatment of Shiga toxigenic Escherichia coli infections and dysentery in humans. *Nat Med* **6**, 265-270.

Paton, A.W., Srimanote, P., Woodrow, M.C., and Paton, J.C. (2001). Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic Escherichia coli strains that are virulent for humans. *Infect Immun* **69**, 6999-7009.

Perissi, V., Staszewski, L.M., McInerney, E.M., Kurokawa, R., Krones, A., Rose, D.W., Lambert, M.H., Milburn, M.V., Glass, C.K., and Rosenfeld, M.G. (1999). Molecular determinants of nuclear receptor-corepressor interaction. *Genes Dev* **13**, 3198-3208.

Perna, N.T., Mayhew, G.F., Posfai, G., Elliott, S., Donnenberg, M.S., Kaper, J.B., and Blattner, F.R. (1998). Molecular evolution of a pathogenicity island from enterohemorrhagic Escherichia coli O157:H7. *Infect Immun* **66**, 3810-3817.

Phillips, N., Hayward, R.D., and Koronakis, V. (2004). Phosphorylation of the enteropathogenic E. coli receptor by the Src-family kinase c-Fyn triggers actin pedestal formation. *Nat Cell Biol* **6**, 618-625.

Pochetti, G., Godio, C., Mitro, N., Caruso, D., Galmozzi, A., Scurati, S., Loiodice, F., Fracchiolla, G., Tortorella, P., Laghezza, A., *et al.* (2007). Insights into the mechanism of partial agonism: crystal structures of the peroxisome proliferator-activated receptor gamma

ligand-binding domain in the complex with two enantiomeric ligands. *J Biol Chem* **282**, 17314-17324.

Pruimboom-Brees, I.M., Morgan, T.W., Ackermann, M.R., Nystrom, E.D., Samuel, J.E., Cornick, N.A., and Moon, H.W. (2000). Cattle lack vascular receptors for Escherichia coli O157:H7 Shiga toxins. *PNAS* **97**, 10325-10329.

Ramegowda, B., and Tesh, V.L. (1996). Differentiation-associated toxin receptor modulation, cytokine production, and sensitivity to Shiga-like toxins in human monocytes and monocytic cell lines. *Infect Immun* **64**, 1173-1180.

Rangwala, S.M., and Lazar, M.A. (2002). The dawn of the SPPARMs? *Sci STKE* **2002**, pe9.

Reid, S.D., Herbelin, C.J., Bumbaugh, A.C., Selander, R.K., and Whittam, T.S. (2000). Parallel evolution of virulence in pathogenic Escherichia coli. *Nature* **406**, 64-67.

Richardson, S.E., Rotman, T.A., Jay, V., Smith, C.R., Becker, L.E., Petric, M., Olivieri, N.F., and Karmali, M.A. (1992). Experimental verocytotoxemia in rabbits. *Infect Immun* **60**, 4154-4167.

Ritchie, J.M., and Waldor, M.K. (2005). The locus of enterocyte effacement-encoded effector proteins all promote enterohemorrhagic Escherichia coli pathogenicity in infant rabbits. *Infect Immun* **73**, 1466-1474.

Robert Koch-Institut (2003). Ein HUS-Ausbruch durch Sorbitol-fermentierende EHEC des Serovars O157:H–: Untersuchungsergebnisse und Lehren für die Surveillance. *Epidemiologisches Bulletin*, 171-178.

Robinson, C.M., Sinclair, J.F., Smith, M.J., and O'Brien, A.D. (2006). Shiga toxin of enterohemorrhagic Escherichia coli type O157:H7 promotes intestinal colonization. *PNAS* **103**, 9667-9672.

Rosen, E.D., Hsu, C.H., Wang, X., Sakai, S., Freeman, M.W., Gonzalez, F.J., and Spiegelman, B.M. (2002). C/EBPalpha induces adipogenesis through PPARgamma: a unified pathway. *Genes Dev* **16**, 22-26.

Rother, R.P., Rollins, S.A., Mojcik, C.F., Brodsky, R.A., and Bell, L. (2007). Discovery and development of the complement inhibitor eculizumab for the treatment of paroxysmal nocturnal hemoglobinuria. *Nat Biotechnol* **25**, 1256-1264.

Rupp, B. (2010). In Biomolecular crystallography (Garland Science), pp. 87-91.

Saarikangas, J., Zhao, H., Pykalainen, A., Laurinmaki, P., Mattila, P.K., Kinnunen, P.K., Butcher, S.J., and Lappalainen, P. (2009). Molecular mechanisms of membrane deformation by I-BAR domain proteins. *Curr Biol* **19**, 95-107.

Sallee, N.A., Rivera, G.M., Dueber, J.E., Vasilescu, D., Mullins, R.D., Mayer, B.J., and Lim, W.A. (2008). The pathogen protein EspF(U) hijacks actin polymerization using mimicry and multivalency. *Nature* **454**, 1005-1008.

Sarraf, P., Mueller, E., Smith, W.M., Wright, H.M., Kum, J.B., Aaltonen, L.A., de la Chapelle, A., Spiegelman, B.M., and Eng, C. (1999). Loss-of-function mutations in PPAR gamma associated with human colon cancer. *Mol Cell* **3**, 799-804.

Sawallisch, C., Berhorster, K., Disanza, A., Mantoani, S., Kintscher, M., Stoenica, L., Dityatev, A., Sieber, S., Kindler, S., Morellini, F.*, et al.* (2009). The insulin receptor substrate of 53 kDa (IRSp53) limits hippocampal synaptic plasticity*. J Biol Chem* **284**, 9225-9236.

Schmidt, H., Beutin, L., and Karch, H. (1995). Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7 strain EDL 933*. Infect Immun* **63**, 1055-1061.

Schmidt, H., Henkel, B., and Karch, H. (1997). A gene cluster closely related to type II secretion pathway operons of gram-negative bacteria is located on the large plasmid of enterohemorrhagic Escherichia coli O157 strains*. FEMS Microbiol Lett* **148**, 265-272.

Schmitz, A.M., Morrison, M.F., Agunwamba, A.O., Nibert, M.L., and Lesser, C.F. (2009). Protein interaction platforms: visualization of interacting proteins in yeast*. Nat Meth* **6**, 500-502.

Schüller, S., Frankel, G., and Phillips, A.D. (2004). Interaction of Shiga toxin from Escherichia coli with human intestinal epithelial cell lines and explants: Stx2 induces epithelial damage in organ culture*. Cell Microbiol* **6**, 289-301.

Schuller, S., Heuschkel, R., Torrente, F., Kaper, J.B., and Phillips, A.D. (2007). Shiga toxin binding in normal and inflamed human intestinal mucosa*. Microbes Infect* **9**, 35-39.

Scita, G., Confalonieri, S., Lappalainen, P., and Suetsugu, S. (2008). IRSp53: crossing the road of membrane and actin dynamics in the formation of membrane protrusions*. Trends Cell Biol* **18**, 52-60.

Sekerkova, G., Loomis, P.A., Changyaleket, B., Zheng, L., Eytan, R., Chen, B., Mugnaini, E., and Bartles, J.R. (2003). Novel espin actin-bundling proteins are localized to Purkinje cell dendritic spines and bind the Src homology 3 adapter protein insulin receptor substrate p53*. J Neurosci* **23**, 1310-1319.

Shao, D., Rangwala, S.M., Bailey, S.T., Krakow, S.L., Reginato, M.J., and Lazar, M.A. (1998). Interdomain communication regulating ligand binding by PPAR-gamma*. Nature* **396**, 377-380.

Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A., and Greene, G.L. (1998). The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen*. Cell* **95**, 927-937.

Shimizu, T., Kawakami, S., Sato, T., Sasaki, T., Higashide, M., Hamabata, T., Ohta, T., and Noda, M. (2007). The serine 31 residue of the B subunit of Shiga toxin 2 is essential for secretion in enterohemorrhagic Escherichia coli*. Infect Immun* **75**, 2189-2200.

Shimizu, T., Ohta, Y., and Noda, M. (2009). Shiga toxin 2 is specifically released from bacterial cells by two different mechanisms*. Infect Immun* **77**, 2813-2823.

Siegler, R.L. (1995). The hemolytic uremic syndrome*. Pediatr Clin North Am* **42**, 1505-1529.

Sinclair, J.F., and O'Brien, A.D. (2002). Cell surface-localized nucleolin is a eukaryotic receptor for the adhesin intimin-gamma of enterohemorrhagic Escherichia coli O157:H7*. J Biol Chem* **277**, 2876-2885.

Smith, K., Humphreys, D., Hume, P.J., and Koronakis, V. (2010). Enteropathogenic Escherichia coli Recruits the Cellular Inositol Phosphatase SHIP2 to Regulate Actin-Pedestal Formation*. Cell Host Microbe* **7**, 13-24.

Soltyk, A.M., MacKenzie, C.R., Wolski, V.M., Hirama, T., Kitov, P.I., Bundle, D.R., and Brunton, J.L. (2002). A mutational analysis of the globotriaosylceramide-binding sites of verotoxin VT1. *J Biol Chem* **277**, 5351-5359.

Stahl, A.L., Sartz, L., Nelsson, A., Bekassy, Z.D., and Karpman, D. (2009). Shiga toxin and lipopolysaccharide induce platelet-leukocyte aggregates and tissue factor release, a thrombotic mechanism in hemolytic uremic syndrome. *PLoS ONE* **4**, e6990.

Stein, P.E., Boodhoo, A., Tyrrell, G.J., Brunton, J.L., and Read, R.J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from E. coli. *Nature* **355**, 748-750.

Suetsugu, S., Murayama, K., Sakamoto, A., Hanawa-Suetsugu, K., Seto, A., Oikawa, T., Mishima, C., Shirouzu, M., Takenawa, T., and Yokoyama, S. (2006). The RAC binding domain/IRSp53-MIM homology domain of IRSp53 induces RAC-dependent membrane deformation. *J Biol Chem* **281**, 35347-35358.

Sugatani, J., Igarashi, T., Munakata, M., Komiyama, Y., Takahashi, H., Komiyama, N., Maeda, T., Takeda, T., and Miwa, M. (2000). Activation of coagulation in C57BL/6 mice given verotoxin 2 (VT2) and the effect of co-administration of LPS with VT2. *Thromb Res* **100**, 61-72.

Swimm, A., Bommarius, B., Reeves, P., Sherman, M., and Kalman, D. (2004). Complex kinase requirements for EPEC pedestal formation. *Nat Cell Biol* **6**, 795-795.

Taylor, F.B., Jr., Tesh, V.L., DeBault, L., Li, A., Chang, A.C., Kosanke, S.D., Pysher, T.J., and Siegler, R.L. (1999). Characterization of the baboon responses to Shiga-like toxin: descriptive study of a new primate model of toxic responses to Stx-1. *Am J Pathol* **154**, 1285-1299.

Thierry, T., Richard, D.H., Jeyanthy, E., John, M.L., and Vassilis, K. (2004). Self-association of EPEC intimin mediated by the beta-barrel-containing anchor domain: a role in clustering of the Tir receptor. *Mol Microbiol* **51**, 73-87.

Tobe, T., Beatson, S.A., Taniguchi, H., Abe, H., Bailey, C.M., Fivian, A., Younis, R., Matthews, S., Marches, O., Frankel, G.*, et al.* (2006). An extensive repertoire of type III secretion effectors in Escherichia coli O157 and the role of lambdoid phages in their dissemination. *PNAS* **103**, 14941-14946.

Tontonoz, P., Graves, R.A., Budavari, A.I., Erdjument-Bromage, H., Lui, M., Hu, E., Tempst, P., and Spiegelman, B.M. (1994). Adipocyte-specific transcription factor ARF6 is a heterodimeric complex of two nuclear hormone receptors, PPAR gamma and RXR alpha. *Nucleic Acids Res* **22**, 5628-5634.

Tontonoz, P., and Spiegelman, B.M. (2008). Fat and beyond: the diverse biology of PPARgamma. *Annu Rev Biochem* **77**, 289-312.

Toshima, H., Yoshimura, A., Arikawa, K., Hidaka, A., Ogasawara, J., Hase, A., Masaki, H., and Nishikawa, Y. (2007). Enhancement of Shiga toxin production in enterohemorrhagic Escherichia coli serotype O157:H7 by DNase colicins. *Appl Environ Microbiol* **73**, 7582-7588.

Vagin, A., and Teplyakov, A. (1997). MOLREP: An automated programm for molecular replacement. *J Appl Crystallogr* **30**, 1022-1025.

Vingadassalom, D., Kazlauskas, A., Skehan, B., Cheng, H.C., Magoun, L., Robbins, D., Rosen, M.K., Saksela, K., and Leong, J.M. (2009). Insulin receptor tyrosine kinase substrate

links the E. coli O157:H7 actin assembly effectors Tir and $EspF_U$ during pedestal formation. *PNAS* **106**, 6754-6759.

Waku, T., Shiraki, T., Oyama, T., Fujimoto, Y., Maebara, K., Kamiya, N., Jingami, H., and Morikawa, K. (2009). Structural insight into PPARgamma activation through covalent modification with endogenous fatty acids. *J Mol Biol* **385**, 188-199.

Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-134.

Weinstein, D.L., Jackson, M.P., Perera, L.P., Holmes, R.K., and O'Brien, A.D. (1989). In vivo formation of hybrid toxins comprising Shiga toxin and the Shiga-like toxins and role of the B subunit in localization and cytotoxic activity. *Infect Immun* **57**, 3743-3750.

Weidner, C., de Groot, J.C., Prasad, A., Freiwald, A., Quedenau, C., Kliem, M., Kodelja, V., Han, C., Giegold, S., Baumann, M., Klebl, B., *et al.* (submitted). Amorfrutins are potent anti-diabetic dietary natural products.

Weiss, S.M., Ladwein, M., Schmidt, D., Ehinger, J., Lommel, S., Städing, K., Beutling, U., Disanza, A., Frank, R., Jänsch, L., *et al.* (2009). IRSp53 links the enterohemorrhagic E. coli effectors Tir and $EspF_U$ for actin pedestal formation. *Cell Host Microbe* **5**, 244-258.

Welinder-Olsson, C., and Kaijser, B. (2005). Enterohemorrhagic Escherichia coli (EHEC). *Scand J Infect Dis* **37**, 405-416.

Wilson, D.R., and Finlay, B.B. (1997). The 'Asx-Pro turn' as a local structural motif stabilized by alternative patterns of hydrogen bonds and a consensus-derived model of the sequence Asn-Pro-Asn. *Protein Eng* **10**, 519-529.

Wong, C.S., Jelacic, S., Habeeb, R.L., Watkins, S.L., and Tarr, P.I. (2000). The risk of the hemolytic-uremic syndrome after antibiotic treatment of Escherichia coli O157:H7 infections. *N Engl J Med* **342**, 1930-1936.

Wu, Y., Chin, W.W., Wang, Y., and Burris, T.P. (2003). Ligand and coactivator identity determines the requirement of the charge clamp for coactivation of the peroxisome proliferator-activated receptor gamma. *J Biol Chem* **278**, 8637-8644.

Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M.A., Jordan, D.M., Francetic, O., Kaper, J.B., Puente, J.L., and Giron, J.A. (2007). Intestinal adherence associated with type IV pili of enterohemorrhagic Escherichia coli O157:H7. *J Clin Invest* **117**, 3519-3529.

Yamagishi, A., Masuda, M., Ohki, T., Onishi, H., and Mochizuki, N. (2004). A novel actin bundling/filopodium-forming domain conserved in insulin receptor tyrosine kinase substrate p53 and missing in metastasis protein. *J Biol Chem* **279**, 14929-14936.

Yanagida-Asanuma, E., Asanuma, K., Kim, K., Donnelly, M., Young Choi, H., Hyung Chang, J., Suetsugu, S., Tomino, Y., Takenawa, T., Faul, C., *et al.* (2007). Synaptopodin protects against proteinuria by disrupting Cdc42:IRSp53:Mena signaling complexes in kidney podocytes. *Am J Pathol* **171**, 415-427.

Zhang, X., McDaniel, A.D., Wolf, L.E., Keusch, G.T., Waldor, M.K., and Acheson, D.W. (2000). Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. *J Infect Dis* **181**, 664-670.

Zimmerhackl, L.B., Verweyen, H., Gerber, A., Karch, H., and Brandis, M. (2002). Das hämolytisch-urämische Syndrom. *Deutsches Ärzteblatt* **4**, A196-A203.

4 Accession Numbers

Atomic coordinates have been deposited in the Protein Data Bank http://www.pdb.org with the PDB IDs given in Table 2.7-1.

Table 2.7-1: Accession Numbers.

Complex	Title of the entry	PDB ID
IMD_{IRSp53}:Tir	Crystal structure of the I-BAR domain of IRSp53 (BAIAP2) in complex with an EHEC derived Tir peptide	2YKT
hPPARγ-LBD:amorfrutin_1	Ligand-binding domain of human PPAR gamma in complex with amorfrutin 1	2YFE
hPPARγ-LBD:amorfrutin_2	Ligand-binding domain of human PPAR gamma in complex with amorfrutin 2	4A4V
hPPARγ-LBD:amorfrutin_B	Ligand-binding domain of human PPAR gamma in complex with amorfrutin B	4A4W

Danksagung

Zunächst möchte ich mich herzlich bei Dr. Konrad Büssow für die Aufnahme in seine Projektgruppe, spannende interdisziplinäre Thematiken und eine erstklassige Betreuung bedanken. Sein großes Interesse an meiner Arbeit sowie seine stetige Diskussionsbereitschaft haben entscheidend zum schnellen, effektiven und erfolgreichen Forschen und somit zum Gelingen dieser Arbeit beigetragen. Zudem möchte ich mich auch für die Unterstützung und Überarbeitung unserer Publikation sowie dieser Dissertation und ein sehr angenehmes Arbeitsklima bedanken.

Prof. Dr. Dirk Heinz danke ich vor allem für die Möglichkeit, meine Doktorarbeit in seiner Abteilung anfertigen zu dürfen. Die hervorragenden Arbeitsbedingungen, exzellentes Equipment sowie sein Bestreben, Kooperationen effektiv mit meinen Projekten zu verknüpfen, haben deutlich zur Qualität und Abschluss dieser Arbeit beigetragen. Zudem möchte ich ihm sehr für die Unterstützung bei unserer Veröffentlichung danken, sowie für die Übernahme der Betreuung meines Promotionsvorhabens als Mentor.

Meinem Arbeitsgruppenleiter Dr. Joop van den Heuvel möchte ich ebenfalls für seine Diskussionsbereitschaft, sein großes Interesse am Gelingen dieser Arbeit, seine uneingeschränkte Unterstützung und für viele Tipps und Tricks im Laboralltag danken.

Prof. Dr. Michael Steinert möchte ich herzlich für die bereitwillige Übernahme des Zweitgutachtens danken.

Ebenso möchte ich mich herzlich bei Prof. Dr. Ralf-Rainer Mendel für die Übernahme des Prüfungsvorsitzes bedanken.

Bei Prof. Dr. Theresia Stradal und Kai Schlüter, sowie der restlichen AG Stradal der Uni Münster, möchte ich mich für die erfolgreiche Kooperation, Diskussionen und Hilfestellungen bedanken.

Bei Dr. Sascha Sauer, Dr. Christopher Weidner und Claudia Quedenau der Gruppe für *Nutrigenomics and Gene Regulation* des Max-Planck-Instituts in Berlin möchte ich mich ebenfalls herzlich für eine hervorragende Kooperation bedanken.

Bei Dr. Ronald Frank bedanke ich mich für die bereitwillige Teilnahme an meinem Thesis-Komitee und anregenden Diskussionen.

Dr. Victor Wray danke ich für die sehr hilfreichen Tipps, Anregungen und tatkräftige Unterstützung bei der Formulierung und Gestaltung unserer Publikation.

Bei Dr. Thorsten Lührs bedanke ich mich für die Durchführung der AF4 Experimente.

Bei Dr. Manfred Nimtz, Undine Felgenträger und Anja Meier möchte ich mich für eine Vielzahl von massenspektroskopischen Analysen bedanken.

Bei Yvonne Carius, Jörn Krausze, Stefan Schmelz, Joachim Reichelt, Christian Strube, Matthias Haffke, Lilia Polle, Agnes Zimmer, Steffen Meyer, Thomas Heidler und Nick Quade möchte ich mich herzlich für viele Diskussionen, Hilfestellungen und eine tolle Zeit im Labor, Grafikraum und am Synchrotron bedanken.

Bei unserem Hiwi Adrian Zander möchte ich mich für seine Mitarbeit und sein Engagement im Labor bedanken.

Ich danke Ute Widow, Daniela Gebauer, Nadine Konisch und Sarah Tokarski für die Einarbeitung, eine sehr herzliche Arbeitsatmosphäre und für alles, wofür der Platz hier einfach nicht ausreichen würde.

Ich möchte mich bei allen Mitgliedern der RPEX und der gesamten SB für ein perfektes Arbeitsklima, eine gute Zusammenarbeit und eine sehr schöne und lehrreiche Zeit bedanken.

Ein besonderer Dank gilt zudem Alexander Eberth, Slava Ziegler und Katharina Lecher für ihre sehr hilfreichen Unterstützungen.

Frau Dr. Willeke, Herrn Dr. Weiland und Frau Dr. Härtha-Thoma danke ich für schnelle und richtige Entscheidungen im richtigen Moment, die mich wieder auf die Beine gebracht haben und u. a. auch den Abschluss dieser Arbeit erst ermöglichten.

Abschließend möchte ich mich bei meiner Familie und meiner Freundin Jana bedanken, die mich in jeglicher Hinsicht und zu jeder Zeit unterstützt haben.